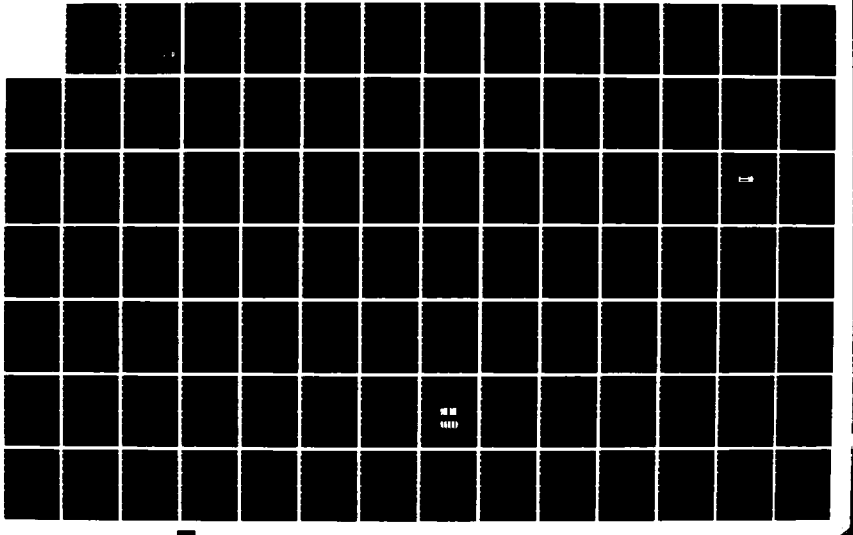
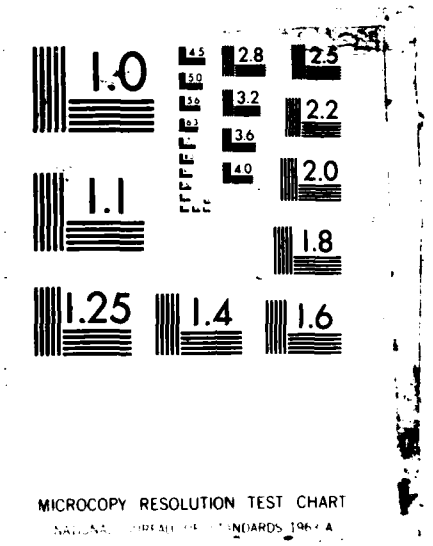
AD-A179 332 EVALUATING THE OPERATIONAL CONTENT OF CONTESTABLE MARKET THEORY(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING 1/2

UNCLASSIFIED P M LUNDY DEC 86 AFIT/GOR/OS/86D-9 F/G 5/3 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

AFIT/GOR/OS/86D-9

EVALUATING THE OPERATIONAL CONTENT OF CONTESTABLE MARKET THEORY

THESIS

Patrick M. Lundy
Second Lieutenant, USAF

AFIT/GOR/OS/86D-9

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A179 332

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): AFIT/GOR/OS/86D-9	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: School of Engineering; 6b. OFFICE SYMBOL (If applicable): AFIT/ENG	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code): Air Force Institute of Technology, Wright-Patterson AFB, Ohio 45433	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification): See Box 19

12. PERSONAL AUTHOR(S): Patrick M. Lundy, B.S., 2d Lt, USAF

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
MS Thesis	FROM _____ TO _____	1986 December	141

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	
05	03		Economic Theory

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Title: Evaluating the Operational Content of Contestable Market Theory

Thesis Chairman: Robert F. Allen
Professor of Economic Analysis

IAW AFR 190-1.
LYNN E. WELAER
Dean for Research and Professional Development
Air Force Institute of Technology (AFIT)
Wright-Patterson AFB OH 45433
7 April 87

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Patrick M. Lundy, 2d Lt, USAF	513-255-3362	AFIT/ENS

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

Abstract

Contestable markets are defined and the properties of these markets are developed. Conventional entry barrier theory is reviewed and contrasted with contestable market theory. The criticisms, experimental evidence, and empirical studies relating to the theory are presented. An unpublished study by Allen (undated) is reviewed and discussed in terms of contestable market theory. Allen's analysis does not support the theory. Considerations in developing and implementing an empirical test of contestable market theory are presented. The theory is not considered useful as a general tool for market analysis.

AFIT/GOR/OS/86D-9

EVALUATING THE OPERATIONAL CONTENT OF CONTESTABLE MARKET THEORY

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Operations Research

Patrick M. Lundy, B.S.

Second Lieutenant, USAF

December 1982

Acknowledgement

I gratefully acknowledge the patience, guidance, and assistance provided by my thesis advisor Dr. Robert Allen.

> "Is my understanding equal to this task, or not? If it is, I apply it to the work as a tool presented to me by Nature. If not, then either I make way - if my duty permits it - for someone more capable of doing the business, or else I do the best I can with the help of some assistant, who will avail himself of my inspiration to achieve what is timely and serviceable for the community. For everything I do, whether by myself or with another, must have as its sole aim the service and harmony of all." (Aurelius, Book 7, Verse 5)

I also thank my reader, Dr. Dennis Quinn.

Table Of Contents

List of Figures

Figure

List of Tables

Abstract

Contestable Markets are defined and properties of these markets are developed. Conventional entry barrier theory is reviewed and contrasted with contestable market theory. The criticisms, experimental evidence, and empirical studies relating to contestable market theory are presented. An unpublished study by Allen (undated) is reviewed and discussed in terms of contestable market theory. Allen's analysis does not support contestable market theory. Considerations in developing and implementing an empirical test are presented. The theory is not considered useful as a general tool for market analysis.

EVALUATING THE OPERATIONAL CONTENT OF CONTESTABLE MARKET THEORY

I. Introduction

In 1968 Harold Demsetz argued that there was not always a need to regulate natural monopolies in order to achieve competitive outcomes. His argument was based on the idea of competition *for* the market rather than within the market; that is, since demand can support only one producer, if that producer earned supernormal profits a different producer would have incentive to enter the market and charge slightly lower profits. To achieve competitive results, two assumptions were needed in his development:

i) The inputs needed for production would be avialable to any firm at the market price; and

ii) The cost of collusion amoung the producers is prohibitively high.

This is the foundation of contestable market theory.

Briefly, a contestable market, as defined in the current literature (Baumol et al., 1982a:5), is a market where entry and exit is unrestricted and costless, where the incumbent firm is not expected to adjust its price in response to entry, and where the number of potential entrants is large enough to prevent collusion. These

markets, then, are vulnerable to costless hit-and-run entry.

As a consequence of hit-and-run entry, these markets are forced to act as if they were competitive, regardless of the number of firms actually in the market. If any firm in a contestable market is earning a positive economic profit, then another firm can come into the market, offer a slightly lower price, and force the original firm out of the market. If the incumbent firm does issue a price response, then the firm which entered can then leave the market, incurring no loss in the market foray. Thus, to deter entry, no firm will price at greater than average cost even in the short-run.

The inability of a firm to earn a positive profit, even in the short-run, in a contestable market is the key feature of these markets. As a consequence, any multifirm contestable market will satisfy the conditions for a "first best" Pareto optimum, and a natural monoply will yield a "second best" Pareto optimum (that is, welfare will be maximized subject to the constraint that the firm earns a nonnegative profit). Thus, contestable markets have desireable consequences, consequences which differ dramatically from those of conventional industry structure theory.

Contestable market theory has not been universally accepted though. It has been attacked primarily on two fronts: its robustness and its applicability. The assumptions, which can be relaxed a bit, are considered by some economists as unrealistic and counterfactual. Real world

markets may exist which satisfy the assumptions, they say, but they are certainly not the norm. Contestable market theory is just a "special case" analysis, not the general theory it is alleged to be by its supporters.

Concurrent with the above criticisms are criticisms of the theory's robustness vis-a-vis small departures from its assumptions. If the theory cannot be applied exactly, then does the theory fall apart under situations which are close to the assumed conditions? In some of its assumptions the theory is robust, with the expected outcomes moving smoothly as small deviations from the assumptions are introduced. However, small deviations from other assumptions give drastically different outcomes.

With such criticisms, why study contestable markets? First, they do provide a benchmark against which the performance of real world markets may be measured. In this role, contestable markets serve much the same purpose as perfectly competitive markets, providing a useful approximation to many markets. Secondly, contestable market theory offers new prospects for market regulation. If steps can be take to make monopoly markets contestable, then price controls would not be needed, the market would enforce competitive pricing. It was this potential regulatory role that, in fact, spurred the development of the theory.

Problem Statement

The purpose of this research is to evaluate the operational content of contestable market theory. That is, to determine the extent to which the theory is testable, especially as pertains to the implications and predictions of the theory.

Motivation

Contestable market theory has been recently put forward with great claims by its authors as to its potential (Baumol et al., 1982a:xiii; Baumol, 1982b:2). But, as already mentioned, it has been received with skepticism by many economists. One point has been noted by both proponents and critics alike: the need for empirical evidence (Baumol, 1982b:14; Soence, 1983:982;579-584). This effort will make a step in that direction by clarifying how contestable market theory relates to conventional competitive and entry barrier theory, and by examining how and to what extent the theory is testable.

Approach

The problem will be attacked in three phases. The first phase will lay out the theory as it has been developed. The recent developments and criticisms of the theory will be incorperated, and, where necessary, the theory will be expanded to cover different situations.

The second phase will evaluate the evidence supporting the theory. The empirical and experimental data, which is

limited, will be presented. The evidence and the criticisms of it will be discussed to evaluate its worth.

Using this background, the third phase will formulate and discuss a model which might be used to evaluate the operational content of the theory.

Goal

The goal of this research is to gain a clear understanding of the operational content of contestable market theory. Once this understanding is obtained, an empir- ical test can be developed which might be used to evaluate the theory. The next chapter presents a summary development of contestable market theory.

Overview

Briefly, the content of the following chapters is:

Chapter II: The assumptions and results of contestable market theory are presented.

Chapter III: Conventional entry barrier theory is reviewed and contrasted with contestable market theory.

Chapter IV: The process for evaluating the operational content of contestable market theory is developed.

Chapter V: The existing empirical and experimental studies of contestable market theory are discussed.

Chapter VI: An empirical test which might be used to study contestable market theory is developed and factors which must be considered in applying the test are presented.

II. Perfectly Contestable Market Theory

Contestable market theory is an emerging body of industrial structure theory which seeks to develop the properties and consequences of perfectly contestable markets and ultimately to apply the results to the analysis of real world markets. The starting point of the theory is the idea of a perfectly contestable market. As defined by Baumol (Baumol et al., 1982a:5), a perfectly contestable market is accessible to potential entrants and has two additional characteristics:

i) a potential entrant can serve the same market demands as an incumbent firm, and

ii) a potential entrant can evaluate the profitability of entry using the incumbent's pre-entry prices.

A further implicit assumption of contestable market theory is that there are a "large number" of potential entrants for any such market. Here, the large number of firms is any number sufficient to insure that there is no possibility of all those firms forming a cartel should they all enter the market. The rationale for this requirement is first given in the work of Harold Demsetz (1968:60). Consider a market controlled by a small group of firms in the form of a cartel. This cartel, if strong enough, can earn monopoly profits through restriction of supply. If a new firm enters the market, the cartel can bribe the new firm into joining

the cartel by offering that firm a share of the monopoly profits, which will be greater than the competitive rate that the firm could earn on its own. Thus the cartel retains some of its power to restrict supply and earn monopoly profits at the price of reduced profits to each firm. As more and more new firms are drawn into the market by the supernormal profits, each firm's share of the profits will continue to decrease, eventually becoming insignificant. Then the cartel can no longer bribe new firms to join, and it may even have trouble keeping its original members. This, then, is what is meant by a large number of potential entrants. Thus, the large number requirement for contestable markets functions in much the same way as the large number requirement in traditional perfectly competitive market theory.

To serve the same market demands as the incumbent firms, the productive technology employed by them must be freely available to any entrant and there must be no perceived qualitative differences between the potential entrant's and the incumbent's products. That is, any entrant must be able to produce a perfect substitute for the incumbent's good at the same price. Since the goods produced by the incumbent and the entrant are perfect substitutes, the only possible difference, in the eyes of the consumers, is the price of the good, and consumers will always chose the lower price. Thus, the demand for either the incumbent's or

the entrant's good is perfectly elastic relative to the price of the other.

Entering firms realize that the expansion of output due to their entry into a market will lead to lower prices, and they assume that if they undercut the incumbent's price, they can sell as much of their good as the market demands at the price they are offering. Thus, the entering firms assume that there will be no price response by the incumbent firms. This is the essence of the second characteristic of contestable markets: the lack of strategic pricing responses by the incumbent firms allows entering firms to determine the profitability of entry based on the incumbent's pre-entry prices and his own offered prices.

An equivalent way of defining a perfectly contestable market is to require absolutely free entry and costless exit (Baumol, 1982b:3). In this context, free entry is taken to mean that there are no entry barriers as defined by Stigler; there are no costs which an entrant must bear which are not borne by an incumbent firm (Stigler, 1968:67). Thus, free entry requires that an entrant have free access to the same production technology as the incumbents, since if this were not so the potential entrant's production costs would be greater than those of the incumbent, which is, in effect, an entry barrier. Similarly, if the incumbent's product were perceived to be better than the entrant's, the entrant would have to incur more advertising costs or increase the quality (and hence the costs) of his product to produce a perfect

substitute for the entrant's good. Thus, free entry alone is sufficient to guarantee the first characteristic of contestable markets (the free availability of technology and the perfect substitutability of the goods).

Free exit is needed to deal with the second characteristic (lack of pricing responses by the incumbent). Free exit, in this context, not only requires that firms are able to leave the market without restriction, but also that, in leaving, a firm can recoup any costs incurred during its entry into the market (Baumol, 1982b:3). With all entry costs, capital investment in particular, recoverable, a potential entrant does not risk anything by entering the market; if the incumbent cuts prices, the entrant can exit the market and have lost no money. Potential entrants can thus evaluate the profitability of entry based on the incumbent's pre-entry prices, knowing that if the incumbent does change its price, the entrant can leave the market with no loss.

Defining contestable markets in terms of free entry and exit highlights the crucial feature of these markets: vulnerability to hit-and-run entry. Should any opportunity for profit materialize, a potential entrant can enter the market, realize a positive profit, and leave the market if the established firms issue a pricing response, incurring no net costs in the processes of entry and exit. Intuitively, this leads to several desirable consequences. First, no perfectly contestable market can yield more than a normal rate of profit, regardless of the number of firms in the

market. Also, a contestable market can allow no inefficiency in production, input allocation, or industry organization. Any inefficiency would represent an opportunity for positive profits to a group of firms which did not have those inefficiencies, which could then enter the market and displace the established firms. These results and others will be developed more rigorously later in this chapter.

These results are also obtained under perfect competition. This is because any perfectly competitive market is, in fact, perfectly contestable. A perfectly competitive market hypothesizes a homogeneous good (and hence an incumbent's and an entrant's good will be perfect substitutes), freely available production technology, and constancy of price (no incumbent firm will initiate a pricing response). However, the results of contestable market theory are applicable to any contestable market, regardless of the number of firms actually in the market, as will be made explicit later. It is only the ability to enter and exit costlessly that is required. The ability of firms to enter and exit a market freely should not be taken to imply that considerable entry and exit will occur. Contestable market theory does not predict any actual entry or exit. In fact, one of the principal ideas of contestable market theory will be the study of markets with sustainable configurations, which do not permit profitable entry. These ideas will be defined and developed later.

Notation and Explanatory Material

The following points pertain to all discussions from this point onward unless specifically noted otherwise:

1. A firm's output is a (n x 1) vector, where n is the number of different products produced by the firm. A firm's output is denoted by $\underset{\sim}{x}$. Similarly, price is a (n x 1) vector denoted by $\underset{\sim}{p}$.
2. Lowercase or numeral superscripts refer to the vector quantity for that particular firm; $\underset{\sim}{y}^i$ is the output vector for firm i. An uppercase superscript "E" denotes the vector quantity corresponding to an entering firm, and an uppercase "I" corresponds to an incumbent firm.
3. $C(\underset{\sim}{y})$ denotes the least costly way for a firm to produce output $\underset{\sim}{y}$.
4. $Q(\underset{\sim}{p})$ is the market demand function. It denotes the quantity of output demanded by the market at a given price.
5. An industry configuration specifies the number of firms active in the industry, the output of each firm, and the price vector for the industry. It can be written in vector form as $(m,\underset{\sim}{y}^1,\ldots,\underset{\sim}{y}^m,\underset{\sim}{p})$, where m is the number of firms in the industry with positive output (Spence, 1983:982).
6. In the case of a multiproduct firm, all the markets within which the firm participates are contestable. Unless noted otherwise, all discussion applies to

all industries, regardless of the number of products involved.

In the multiproduct case, the use of the word industry needs some clarification. The boundries of an industry may be defined either in terms of production or in terms of consumption. That is, goods closely related in consumption (e.g., pencils and pens) or goods closely related in production (e.g., airplane and missile engines) may be the basis of inclusion to an industry. The definition used may have some effect on results since, in the one case, there is an implied relation among the production and cost functions, and in the other case there is a demand relation. In most of the work, though, the difference is not significant, and the term industry can be considered to be used in either context (Baumol et al., 1982a:113).

Barriers to Entry

Fixed and Sunk Costs. In the definition of contestable markets, the necessity of free entry and exit was given as a salient characteristic of these markets. At the time there was no discussion as to what characteristics of an industry or market could impede entry or exit. This section will address that issue.

First, two terms must be clearly defined: sunk costs and fixed costs. Sunk costs are costs that cannot be eliminated in the short- or intermediate- run even if production is stopped altogether, but which can be eliminated in the

long-run by reducing or discontinueing production, that is, by exiting the market (Baumol et al., 1982a: 280). Fixed costs, on the other hand, cannot be eliminated, even in the long-run, as long as production is still positive (Baumol et al., 1982a:280). Formally, sunk costs can be defined as follows:

SUNK COSTS:

Define $C(\underset{\sim}{y},s)$ to be the short-run cost function applicable over the next s time periods. Then $K(s)$ is the sunk cost for s periods if:

$$C(\underset{\sim}{y},s) = K(s) + G(\underset{\sim}{y},s) \qquad (1)$$

where

$$G(0,s) = 0 \qquad (2)$$

and

$$\lim\{ K(s): s \to \infty \} = 0 \qquad (3)$$

(that is, sunk costs are zero in the long-run) (Coursey et al., 1982a:70).

And, for fixed costs:

FIXED COSTS:

Define $C_L(\underset{\sim}{y})$ to be the long-run total cost function. Then the long-run fixed cost is F if:

$$C_L(\underset{\sim}{y}) = B(y)F + V(\underset{\sim}{y}) \qquad (4)$$

where

$$\lim\{ V(\underset{\sim}{y}): \underset{\sim}{y} \to 0 \} = 0 \quad , \qquad (5)$$

$V(\underset{\sim}{y})$ is nondecreasing in $\underset{\sim}{y}$,

and

$$B(\underset{\sim}{y}) = 1 \text{ if } \underset{\sim}{y} > 0 \qquad (6)$$

$$= 0 \text{ if } \underset{\sim}{y} = 0$$

(Baumol et al., 1982a:280).

To emphasize the difference between these two concepts, consider the following examples (Baumol et al., 1982a: 281-282). Auto manufacturing is generally a capital intensive industry. But, using the above definitions, these capital costs are sunk costs, but not fixed. If the car market declined drastically, the car manufacturers, if they decided to remain in the industry, eventually would begin to manufacture the cars by hand, avoiding the high replacement costs of the assembly machinery. Thus, the cost of the assembly equipment is sunk for the life of the equipment, since, once purchased, the costs cannot be eliminated in the short-run by ceasing production, but they can be elimninated in the long-run while still maintaining positive production.

For an example of fixed costs that are not sunk, consider the market for airline travel between a pair of cities. If passengers are going to be transported along this route (that is, if production is positive), at least a single-engine plane is required. This cost is fixed (up to the capacity of the plane), but it is not sunk. The plane is needed to maintain positive production, but it can easily be diverted to another market, and hence, in the short-run, the cost can be avoided.

For production levels up to its capacity, the cost of any capital required for production will be fixed if it can be easily and costlessly resold or converted to other uses. First, the capital must be required for efficient production since its cost is included in the long-run cost function with positive production (by the definition of fixed costs). Second, as long as production is within the capital's capacity, no additional capital will be needed for production. Thus, the capital cost will be fixed for production levels less than its capacity. Conversely, if production is considered to an infinite level of output, then every factor would be variable. Lastly, given capital meeting the above requirements, if it can be costlessly be converted to other uses, that is, if it is highly fungible capital, its costs will be mostly fixed. Since such capital can quickly be converted to other uses or sold, its costs can be recovered in the short-run, and, hence, it is not sunk.

Highly fungible capital includes such things as simple capital which can be sold at little or no loss for scrap (I-beams for instance), multipurpose or generic capital that can be used in many different industries (deliveriy trucks, hammers, drill presses, etc.), and "capital on wheels", such as airplanes, delivery trucks, and trains on a specific route, if that specific route is considered to be the appropriate market. In the first two examples, there is generally a significant resale market for the machinery, so that most of the costs can be recovered. In the last case, there

are usually other markets, that is, routes, to which the capital can be shifted without cost. Also, the cost of short-term rental equipment is a fixed cost to the firm using the equipment, since the cost can be avioded in the short-run.

The cost of highly specialized capital, on the other hand, is generally a sunk cost. There will probably not be a significant resale market for equipment used in micro-surgery or the manufacture of silicon chips, especially if the firms leave the markets because they are crowded out. The difference between this case and the highly fungible goods is that the highly fungible goods can be resold to producers in other markets; a forklift used by a firm manufacturing diapers is just as useful to a firm manufacturing shotguns, while a scapel used in microsurgery is not going to be of much use to any other firm. Capital which is leased on a long-term basis will also incur sunk costs (if it cannot be put to other uses). Similarly, licensing fees, advertising expenditures, investment in research and development, and technical training costs are all sunk costs to the degree that they are product specific.

Fixed Costs, Sunk Costs, and Entry Barriers. Before some propositions on entry barriers can be proven, the term must be clearly defined. The definition that is used is: "An entry barrier is anything that requires an expenditure by a new entrant into an industry, but imposes no equivalent cost upon an incumbent" (Baumol et al., 1982a:282). This

definition corresponds closely with the definition given by Stigler (1968:67). Thus, as will be demonstrated, sunk costs can be a barrier to entry, while neither fixed costs nor economies of scale constitute entry barriers.

PROPOSITION 1: (Baumol et al., 1982a:289-290)

Fixed costs are not entry barriers.

Proof:

Fixed costs are not entry barriers simply because they affect both incumbent and entering firms in the same manner. In the decision to produce or not, both incumbents and entrants must include fixed costs in their calculations. Q.E.D.

This proposition does not imply that high levels of fixed costs will not deter additional firms from entering an industry, but rather that if such entry deterrence occurs, the market must nonetheless display the welfare properties associated with contestable markets (which will be developed later). Indeed, high levels of fixed costs can guarantee that a single firm can find prices which will keep additional firms from entering the market (by ensuring that entry cannot be profitable). When fixed costs are high, firms with larger outputs are able to spread this fixed cost over more units, and, if the fixed costs are sufficiently high, this lower average fixed cost may offset higher average variable costs to allow for lower average total cost at

high levels of output (Baumol et al.,1982a:286). Thus, with high fixed costs, the large-scale producer will be able to offer goods at a lower unit price. The next proposition highlights the distinction between fixed and sunk costs.

PROPOSITION 2: (Baumol et al., 1982a:290-291)

The presence of sunk costs can be a barrier to entry.

Proof:

The possibility that sunk costs can pose a barrier to entry arises from the fact that they should be treated differently in the profit calculations of incumbent and entering firms. Intuitively, entering firms must include the possibility that a portion of the sunk costs might be lost in a market foray, while the incumbent firm has already put the capital at risk, and need not include it in his profit calculations.

Consider an industry which requires that a cost of K be sunk, say in capital, for a given period. Over this period, the incumbent firm will continue production as long as his expected profit, $E(\pi^I)$, is nonnegative; i.e., if $E(\pi^I) \geq 0$.

The entrant, on the other hand, will enter the market if his expected profit from operations in the market, $E(\pi^E)$, is not less than the expected costs of getting into and out of the market, $E(C^E)$; that is, if $E(\pi^E) \geq E(C^E)$. If the entrant remains in the market for the full life of the sunk cost, then he can fully depreciate the capital, and the cost

of entering and exiting the market is zero. If, however, the entrant is forced out of the market before the capital can be fully depreciated, the undepreciated capital is lost if it cannot be resold. Thus, if there is a nonzero probability that the entrant will be forced out of the market without recovering capital costs, there is an expected cost of entering and exiting the market for a potential entrant. This expected cost is the source of the entry barrier, since there is no equivalent cost forced on the incumbent firm.

Note that if a potential entrant can be guaranteed that he will remain in the market for the full life of the sunk cost, then the expected cost of entering and exiting the market is zero. In such a case sunk costs will not raise entry barriers. Q.E.D.

In the above proof, the mechanism which forced the entrant out of the market was not specified. The exact mechanism is inconsequential; it could occur through a natural constriction of the market, through pricing responses initiated by incumbent firms, or through any other mechanism. This proposition is usually applied to justify the prohibition of pricing responses by incumbent firms (Baumol et al., 1982a:300), since that is the easiest source of uncertainty to regulate.

At this point it is also appropriate to briefly discuss the roles of exit barriers. All the work that has been done so far has concentrated on the role of entry barriers, but

in the definition of contestable markets the absence of exit barriers was also required. This emphasis is found in most of the literature on the theory. Shepherd does note that in contestable markets, exit barriers are important in deterring entry only if they are higher than the entry barriers since, if entry barriers are higher, the firms will not enter anyway. Thus, the past literature on entry barriers has merely assumed that the entry barriers are larger than the exit barriers (Shepherd, 1984:578-579).

Some clarification on this issue is needed. First comes the question of whether the two can indeed be distinguished. If there is a cost of exiting the market, this cost is going to have to be included in the profitability calculations for any firm planning to enter the market (if there is a nonzero probability that the firm will leave the market). Thus, it will influence the decision of the firm to enter the market. The only distinction between an entry and exit barrier is that an entry barrier, by definition, is a cost that only the entrant must pay, whereas an exit cost will have to be considered in the production decisions of both the incumbent and the entering firms. The next question which arises is whether both of these barriers must be considered simultaneously in an entrant's production decision, or whether only the higher of the two values is important, as Shepherd suggests. Consider a firm deciding to enter a market. In its calculations, the firm expects to earn $50 dollars in profit and incur entry costs of $45.

The firm would make a net profit of $5 if exit were costless. According to Shepherd, the firm would decide to enter the market as long as the cost of exiting the market is less than $45. But, if exit costs are greater than $5, then the firm will not earn a nonnegative profit, and entry into the market would not be logical. Thus, entry and exit costs are equally important in an entrants decision to enter a market.

Exit costs also have another distorting effect. In the presence of exit costs, a firm would be willing to tolerate negative profits while in the market if the expected loss due to market operations is less than the cost of leaving the market. Strategic pricing responses by an incumbent firm thus become more costly to that firm. It is clear that the presence of exit costs makes the entry and exit decisions more complicated, and the whole issue needs to be explored in more detail.

Scale economies also need not be entry barriers, even though, as in the case of high fixed costs, scale economies over a large range of output can deter additional firms from entering a market. The cost differences due to scale economies do not have to be paid by the entering firm. To avoid paying this cost difference, all the firm needs to do is to produce at the same level as the incumbent firm. Thus, sunk costs can be entry barriers, while neither fixed costs nor scale economies are. There are other collateral issues which may increase the likelihood of entry barriers.

Primarily, in the case of high fixed costs, the means by which firms can obtain the money to cover these costs is never addressed by Baumol et al. High fixed costs increase the likelihood that entering firms would have to borrow money, which, assuming they are able to obtain a loan, would impose an interest cost on them, and hence raise and entry barrier. This same argument would apply to the case of high sunk costs, even if the sinking of the costs did not raise a barrier itself, as noted in Proposition 2. Chapter Three will compare contestrable market theory and conventional entry barrier theory.

Time-Lag Contestability

From Proposition 2 we see that even in the presence of sunk costs, a market can still be contestable if the incumbent firms are prohibited from responding to entry by strategic price responses. This leads to an investigation of contestable markets in terms of time lags. In the market foray, there are three important time lags: the entry lag faced by the entering firm, the price-adjustment lag faced by the incumbent firm, and the exit lag (Dixit, 1982:15-16; Schwartz and Reynolds, 1983: 488-489). The entry lag is the amount of time between a potential entrants recognition of a profitable entry opportunity and the firms establishment in the market. Similarly, the exit lag is the amount of time required by a firm to liquidate all of its assets and leave the market. Both of these lags can be affected by the

amount of capital required, the availability of that captial, and the nature of the capital (Capital which is highly fungible will entail a shorter exit lag). The time needed by an incumbent firm to calculate and implement a price change is the price- adjustment lag. This time lag also includes the time necessary to gain consumer acceptance of the price change.

An alternative condition for contestabitity is then:

PROPOSITION 3:

In a contestable market it is necessary that the entry lag be finite, and either:

1) the price-adjustment lag is greater than the entry lag if the exit lag is finite, or
2) the price-adjustment lag is infinite if the exit lag is infinite.

Proof:

In this proposition if a time lag is infinite, then the corresponding action never occurs. An infinite price-adjustment lag means that prices cannot be changed by the incumbent firms, and an infinite exit lag means that not all the costs incurred during entry are recoverable, that is, there are unrecoverable sunk costs.

If a market is contestable, then the market must be accessable to entrants. The entrants can freely use the same technology as the incumbents and produce an identical

product. Thus, the entry lag must be finite in a contestable market.

A finite exit lag implies that there are no unrecoverable costs. In this case, if the price-adjustment lag were not greater than the entry lag the incumbent firms could wait until entry has actually begun and then adjust prices. The entering firms thus could not evaluate accurately the profitability of entry based on pre-entry prices. A longer price-adjustment lag will guarantee that prices will not change by the time a firm establishes itself in the market.

On the other hand, if there is an infinite exit lag, then there are costs which cannot be recovered upon leaving the market. Thus, according to Proposition 2, the incumbent firms must not be allowed to adjust prices; the price-adjustment lag must be infinite. Q.E.D.

This proposition expands the applicability of contestable market theory. In the definition of contestable markets, no pricing response was allowed by the incumbent. But this proposition allows that condition to be relaxed. All that is now required is that the incumbent be slower moving, in a sense, than an entrant. For a contestable market, then, it is necessary that the price-adjustment lag is the longest of the three lags (Dixit, 1982:16) and that all the lags be sufficiently short in duration. In this context, what constitutes a sufficiently short period of

time is a subjective judgement: Consider an extreme example of a market with entry and exit lags of 100 years each and unadjustable prices. It hardly seems applicable to call the market contestable, even though, by the definitions, it is.

Sustainability and Equilibrium in Contestable Markets

Sustainable Configurations. One of the basic concepts in any study of an industry's configuration is whether or not a configuration is feasible. A feasible configuration is one in which the market demand at the given price is being satisfied and every firm in the industry is earning a nonnegative profit.

FEASIBLE INDUSTRY CONFIGURATION:

An industry configuration $(m, \underset{\sim}{y}^1, \ldots, \underset{\sim}{y}^m, \underset{\sim}{p})$ is feasible if:

$$\text{i)} \quad \sum_{i=1}^{m} \underset{\sim}{y}^i = Q(\underset{\sim}{p}) \quad \text{, and} \tag{7}$$

$$\text{ii)} \quad \underset{\sim}{p} \cdot \underset{\sim}{y}^i \geq C(\underset{\sim}{y}^i) \tag{8}$$

for every $i = 1, 2, \ldots, m$

(Spence, 1983:982).

It is clear that feasibility is necessary condition for a configuration to be in equilibrium.

In a contestable market a concept stronger than feasibility will be used in the discussions of equilibria. This is the important concept of sustainability.

SUSTAINABLE INDUSTRY CONFIGURATION:

A feasible industry configuration $(m, \underset{\sim}{y}^1, \ldots, \underset{\sim}{y}^m, \underset{\sim}{p})$ is

sustainable if $\underset{\sim}{p}^{e} \cdot \underset{\sim}{y}^{e} \leq C(\underset{\sim}{y}^{e})$ for all $\underset{\sim}{p}^{e} \leq \underset{\sim}{p}$ and $\underset{\sim}{y}^{e} \leq Q(\underset{\sim}{p}^{e})$ (Baumol et al., 1982a:313).

If a firm wants to enter a market, it has to offer its goods as a price less than the prevailing market price ($\underset{\sim}{p}^{e} \leq \underset{\sim}{p}$), and it won't produce more of the good than the market demands at that price ($\underset{\sim}{y}^{e} \leq Q(\underset{\sim}{p}^{e})$). In a market characterized by a sustainable industry configuration, a firm cannot enter and earn a positve profit ($\underset{\sim}{p}^{e} \cdot \underset{\sim}{y}^{e} \leq C(\underset{\sim}{y}^{e})$). Thus, a sustainable configuration will not present opportunities for a positive profit, and no new firms will enter the market.

In the work that follows, the following definition of equilibrium will be adopted: a configuration is in equilibrium if it is feasible and if there exists no vector of outputs which can earn a positive profit at current prices. In other words, no firm is losing money and at the current price no firm can make a positive profit at any output level (Spence, 1983:983). A significant result can now be developed: A contestable market in equilibrium must involve a sustainable configuration.

PROPOSITION 4: (Baumol et al., 1983:495)

Any industry configuration in equilibrium must be sustainable, but not every sustainable configuration need be in equilibrium.

Proof:

The first part of this proposition follows directly

from the definitions. Heuristically, if a configuration is sustainable, then no profit can be made by decreasing price or quantity, or both. If no profit can be made by adjusting either price or quantity, then no profit can be made by adjusting only quantity. Thus, the set of equilibrium configurations is a subset of the set of sustainable configurations.

Now, a rigourous proof. If a feasible configuration is in equilibrium, then there exists no output vector which will yield positive profits. That is, $\underset{\sim}{p} \cdot \underset{\sim}{y} \leq C(\underset{\sim}{y})$ for all $\underset{\sim}{y}$ at the given $\underset{\sim}{p}$. Thus, $\underset{\sim}{p}^{E} \cdot \underset{\sim}{y}^{E} \leq \underset{\sim}{p} \cdot \underset{\sim}{y}^{E} \leq C(\underset{\sim}{y}^{E})$, and the configuration is sustainable.

To see that a sustainable configuration need not be in equilibrium, consider the market conditions depicted in Figure 1. The industry configuration ($m = 1$, y^{I}, p^{I}) is sustainable because at any price below p^{I}, say p^{E}, the revenue yielded (the indicated shaded region) from the market is not sufficient to cover total cost, and any firm offering output at that price would incur a loss, as indicated. Thus, there exists no positive opportunity, and the market is sustainable. But this configuration is not in equilibrium since a firm could make a positive profit at current prices simply by offering any quantity above y^{I}. Thus, a sustainable configuration need not be in equilibrium. Q.E.D.

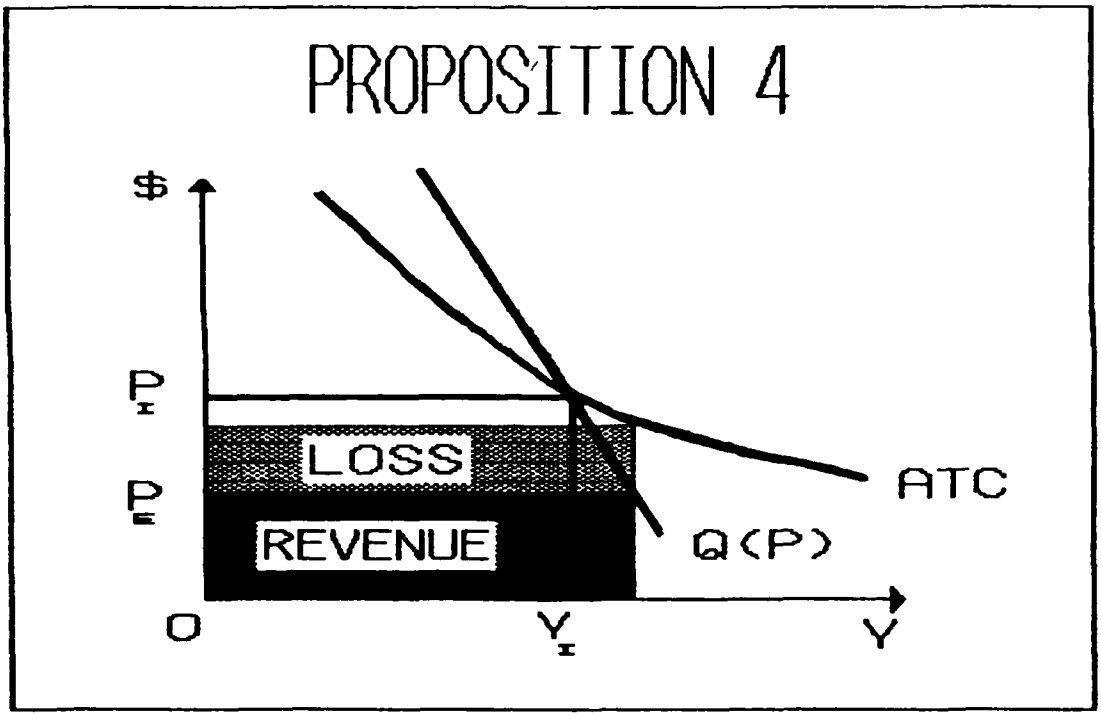


Fig 1. Sustainability and Equilibrium

The above proposition says that any industry configuration which is in long-run competitive equilibrium must be sustainable, regardless of the number of firms in the market. In fact, this proposition is sometimes used to define contestable markets as any market in which sustainability is required for equilibrium (Baumol, 1983:495; Spence, 1983:983), thus highlighting the fact that any perfectly competitive market is also a contestable market. But how does sustainability explicitly relate to contestable markets? The relation is made more clear by considering the effects of free entry and exit. If a given industry is not

sustainable, then there exists a feasible entry plan ($p^e \leq p^x$, $y^e \leq Q(p^e)$) which would yield a positive profit to the entering firm. If there are no costs involved in entry, then this entry plan can be profitably implemented in the market; freedom of entry ensures that any feasible entry plan which yields a positive profit can be successfully implemented in the market.

The linkage between sustainable and equilibrium configurations established in this proposition is crucial. Many of the contestable market propositions begin by hypothesizing a sustainable configuration and then develop its properties. This proposition then guarantees that these propositions will also apply to equilibrium configurations.

Properties of Sustainable Configurations. Now some of the properties of contestable markets will be developed and discussed. In Proposition 4 the linkage between sustainable and equilibrium configurations has already been established. Now, some properties of sustainable, and hence equilibrium, configurations in contestable markets will be developed.

The first result to be discussed is a strong result: In a contestable market, any equilibrium configuration must be an efficient configuration. Here, an efficient configuration is defined as follows:

EFFICIENT INDUSTRY CONFIGURATION:

An industry configuration $(m, \underset{\sim}{y}^1, \ldots, \underset{\sim}{y}^m, \underset{\sim}{p})$ is efficient if the cost of producing the total industry

output with this configuration, $C^x(\underset{\sim}{y})$, is less than the cost of producing the same level of industry output using any other industry configuration (Baumol et al., 1982a:99).

In other words, a configuration is efficient if the same industry output cannot be produced for less. The concept of an efficient industry configuration is quite strong. It implies not only that all the firms in the industry operate efficiently (producing their output at a minimum cost), but also that the resources must be allocated efficiently within the industry.

Now, a proposition:

PROPOSITION 5: (Baumol et al., 1982a:314-316)

In a contestable market, any equilibrium configuration must be efficient.

Proof:

Since any equilibrium configuration must be a sustainable configuration, it is sufficient to show that any sustainable configuration must minimize the total industry cost of producing the output.

Suppose that a sustainable configuration were not efficient. Then there exists a configuration which could produce the same industry output at a lower cost. Since the original configuration was earning a nonnegative profit at the original prices (by assumption it was sustainable, and hence feasible), this new configuration would, in total,

earn positive profits. Thus, there exists a profit opportunity for the efficiently configured group of firms, which contradicts sustainability. Q.E.D.

This proposition provides a crucial link between equilibrium theory in contestable markets and the contemporary theory of cost-minimizing industry structure. Using certain assumptions about the nature of the costs involved in an industry, bounds on the industry cost-minimizing number of firms can be derived. Contestable market theory then says that the number of firms actually in a contestable market must be within these bounds (Baumol et al., 1982a: 316). The actual number of firms in the market will be determined by a combination of the cost and demand conditions characterizing the market.

Proposition 5 provides information about the characteristics of a sustainable configuration in a contestable market. It, however, gives no direct information about the individual firms in the configuration. That is the subject of the next proposition:

PROPOSITION 6: (Baumol et al., 1982a:314)

Any firm in a sustainable configuration in a perfectly contestable market must have the following properties:

(i) The firm must operate efficiently; it must produce its output at minimum cost.

(ii) The firm must earn zero economic profit.

(iii) The firm must select an output vector such that if all outputs are reduced by a factor of k, total costs are reduced by a factor of not more than k ($0 < k < 1$).

(iv) In a multiproduct market, a firm must avoid cross subsidies among any subset of its product set.

Proof:

(i) This is a direct corollary of the industry cost-minimization required by Proposition 5; total industry cost cannot be minimized if any firm in the configuration is not producing its output at minimum cost. This result can also be derived directly by a similar line of reasoning: if a firm is not producing its output efficiently, there is an opportunity for an efficient firm to enter the market at that output level and earn a positive profit. Hence, the original configuration could not have been sustainable.

(ii) If an incumbent firm were earning positive profits for a given price vector, a potential entrant could enter the market with a slightly smaller price vector (that is, no component of the entrant's price vector is greater than the corresponding component of the incumbent's, and at least one component is less than the corresponding component of the incumbent's price vector) and still earn a nonnegative profit. And, since there is an opportunity for a positve profit, the configuration could not have been sustainable.

(iii) Let $\underset{\sim}{y}^z$ represent the output of an incumbent firm in a sustainable configuration, and suppose that an entrant

attempted to enter selling $k\underset{\sim}{y}^z$, $0 < k < 1$, at the same price. Since the original configuration is sustainable, the entrant's marketing plan cannot yield positive profits:

$$C(k\underset{\sim}{y}^z) \geq \underset{\sim}{p}^z \cdot (k\underset{\sim}{y}^z) \qquad (9)$$

Also, an incumbent firm would not operate at a loss, so:

$$\underset{\sim}{p}^z \cdot \underset{\sim}{y}^z \geq C(\underset{\sim}{y}^z) \qquad (10)$$

or,

$$k\underset{\sim}{p}^z \cdot \underset{\sim}{y}^z \geq kC(\underset{\sim}{y}^z) \qquad (11)$$

Combining inequalities (9) and (11) gives:

$$C(k\underset{\sim}{y}^z) \geq kC(\underset{\sim}{y}^z) \quad \text{for all } 0 < k \leq 1 \qquad (12)$$

which is the required inequality (Baumol et al., 1982a:200).

(iv) If an incumbent firm's marketing plan involves cross subsidies, then there is some subset of its product set on which the incumbent is earning a positve profit. An entrant could market only this subset of the incumbent's total product set, offer it at a lower price, and still earn a nonnegative profit, contradicting sustainablility (Brock, 1983:1056).

This result can also be obtained by noting the assumption that all the markets within which a mutliproduct firm operates are contestable then applying part (ii) of this proposition. Since each product earns zero profit, there is no source to provide cross subsidies. Similarly, a firm in a sustainable configuration must be efficient in the

production of each good in its product set by part (i) of this proposition. Q.E.D.

COROLLARY 6(ii)A: (Baumol et al.,1982a:201)

In a sustainable configuration in a contestable market with entry barriers, incumbent firms cannot earn profits greater than the value of the entry barrier.

Proof:

Let $E(\underset{\sim}{y})$ denote the size of the entry barrier. Since an entering firm has to pay the value of the entry barrier in addition to production costs, the total costs to an entering firm are $C(\underset{\sim}{y}^e) + E(\underset{\sim}{y})$. Consider a sustainable configuration with the incumbent firms earning profits greater than the value of the entry barrier; that is:

$$\underset{\sim}{p}^I \cdot \underset{\sim}{y}^I - C(\underset{\sim}{y}^I) > E(\underset{\sim}{y}) \qquad (13)$$

This can be rewritten:

$$\underset{\sim}{p}^I \cdot \underset{\sim}{y}^I - [C(\underset{\sim}{y}^I) + E(\underset{\sim}{y})] > 0 \qquad (14)$$

so that an opportunity for a positive profit exists for an entrant offering the same output as the incumbent ($\underset{\sim}{y}^I = \underset{\sim}{y}^e$) at a slightly lower price. Thus, the original configuration could not have been sustainable.

Note that in a perfectly contestable market there are no entry barriers. The value of the entry barrier is then zero, and this corollary reduces to Proposition 6(ii). Q.E.D.

A corollary similar to this one could be developed for Proposition 6(iv). In this case, if some of the markets within which a multiproduct firm operates have entry barriers then the firm can channel a cross subsidy equal to the height of the entry barrier from any protected market to any other market. That firm would have an advantage in the makrets receiving the subsidies and could offer the goods in those markets at a price less than their cost. The multi-product firm is in effect raising an entry barrier in those markets. Notice how imperfections in one market can propagate to other markets through cross subsidies.

The properties developed in Proposition 6 are independent of the number of firms in the market; these properties will hold in any contestable market regardless of whether the industry consists of one firm, a few firms, or the large number of firms of perfect competition. In this respect contestable market theory unifies industrial organization theory by using one set of assumptions to develop characteristics of competitive, oligopolistic, or monopolistic markets (Baumol, 1982b:2).

So far, we have seen that in contestble markets the price of a good must reflect the minimum cost of producing that good. The pricing system then must reflect technological and organizational efficiency in the industry and within the indivivual firms. All that is required to obtain Pareto optimality of the "first best" variety in the market is some assurance of allocative efficiency (optimally

allocating resources amoung the different goods produced). In a perfectly competitive market this optimality is reflected in the fact that, in equilibrium, price must equal marginal cost. The same result holds for perfectly contestable multifirm markets:

PROPOSITION 7:

In any contestable market with two or more firms, each firm in a sustainable configuration must select an output at which the marginal costs of production are equal to the price of the good (Baumol et al., 1982a:317).

Proof: (Baumol et al., 1982a:26,27; Spence, 1983:983-984)

First, some explanation. In the case of a multiproduct firm, the proposition must hold for each good in the firm's product set. That is, if each market within which the firm is active contains two or more firms, then price will equal marginal cost in each of the markets. To simplify the proof, then, the propostion will be proven for a single output, with the understanding that it must apply to all applicable markets.

This proposition will be proven in two parts. First it will be shown that in any sustainable configuration the marginal cost of a good cannot exceed its price. Then, with the restirction to multifirm sustainable configurations, it will be demonstrated that a good's price cannot be greater than its marginal cost. These two implications then necessitate that the price of a good must equal its marginal cost

in a multifirm sustainable configuration.

Consider an industry with a sustainable configuration. Every firm within the industry is earning zeor profit by Proposition 6(ii). Now, suppose that there was a firm in that configuration for which marginal cost exceeds price. then the cost of producing the last unit of output is greater than the revenue accrued from its sale and, hence, the sale of the last unit of output lowered total profits. So, if an entering firm duplicated that firm's operations except for the last unit of output, the entering fimrs could earn a positive profit. Thus, the original configuration could not have been sustainable.

Similarly, suppose that in a sustainable configuration there existed a firm for which price exceeds marginal cost. This firm's profits can be increased by selling an additional unit of the good. So there is a potential oppornity for earning a positive profit (the original firm had to earn zero profit) by duplicating the original firm's operations and selling one additional unit of output. If there is at least one other firm in the market, that additional unit of output can be sold without a substantial decrease in price, ensuring that profits will increase. Thus, in a multifirm market, if price is greater than marginal cost there exists an opportunity for a positive profit, contradicting sustainability.

Thus, all firms in a sustainable contestable multifirm market must produce at a level for which price equals

marginal cost. Or, if there is only one firm in the market, then the firm must produces at a level where marginal cost price does not exceed. Q.E.D.

As a short corollary to this proposition, every firm in a sustainable configuration must produce at the same marginal cost, since the price will be unique. (Any firm posting a price higher than the minimum will be forced out of the market.) This result is in accordance with the industry cost minimization required by Proposition 5. If all the firms did not have identical marginal costs, total industry costs can be lowered by shifting production from the firms with higher marginal costs to firms with lower marginal costs (Baumol et al., 1982a:26,27).

Natural Monopolies and Contestable Markets

The previous proposition guaranteed a "first best" Pareto optimal resource allocationin any multifirm contestable market. If, however, the market is populated by a single firm, the guarantee is replaced by the guarantee that price is not below marginal cost; all of the results of Proposition 6 still apply. These results are characteristic of a "second best" Pareto optimum; that is, the firm's behavior will maximize welfare subject to the contstraint that the firm does not earn a negative profit. In a single firm market, the firm would lose money if it used marginal cost pricing, since price equals average cost (Proposition

6(ii)) and price (weakly) exceeds marginal costs (Proposition 7). And the price nearest marginal cost that guarantees a nonnegative profit is where price equals average cost. This is a "second best" Ramsey optimum (Baumol et al., 1982a:29).

From Proposition 5 we know that a monopoly can be sustainable in a perfectly contestable market if it is the least costly way of producing the market output. Such industries are called natural monopolies (Baumol et al., 1982a:17). It should be noted that it is the cost and demand structure of the market which will make it a natural monopoly, not the number of firms actually in the market. That is, natural monopoly is a structural not a behavioral characteristic of a market. Since single-firm production is the least costly configuration capable of providing industry output, it is the only sustainable configuration; additional firms may enter the market occassionally, but since they can't earn a profit they will quickly leave.

In a contestable market natural monopolies do not have all of the undesireable characteristics of monopolies in conventional monopolies. Nonetheless, they are not characterized by the same performance as multifirm markets. Thus, we may ask, what conditions will lead to natural monopolies? From the discussion of entry barriers, scale economies over a wide range of outputs levels or high fixed or sunk costs can lead to sustainable natural monopolies. Also, the presence of entry barriers will exacerbate the situation,

allowing artificial monopolies or distorting the "second best" Pareto optimality.

Most of the attention given to contestable market theory in the literature has been focussed on its natural monopoly analysis. This is partially because it is in this area that its results differ most dramatically from conventional analysis. Even natural monopolies in contestable markets must operate efficiently and earn zero profit. Natural monopolies also provide the best market structure within which to test contestable market theory. In large markets contestable market theory is identical with competitive theory; the competitive results are identical with the contestable market results since the competitive assumptions are a special case of the contestable market theory assumptions. But in the case of natural monopolies, the contestable market results are different from conventional results, especially in the presence of significant economies of scale (that is, when the average total cost curve is steep at outputs less than the minimal optimal scale). Under certain assumptions about the incumbent's behavior in the face of entry, significant scale economies do permit the incumbent to earn positive economic profits. But in contestable natural monoplies the incumbent firm cannot earn a positive economic profit, no matter how significant the scale economies. The next chapter will discuss conventional entry barrier theory and contestable market theory, particularly addressing this issue.

Summary

This, then, concludes the overview of contestable market theory. Not every aspect of the theory has been covered; some of the more subtle aspects of the theory, particularly those dealing with the existence of sustainable configurations, have been set aside to better emphasize the principals presented. The second appendix will give a brief discussion of some issues involved with the existence of sustainable configurations in the scalar output case; the multiproduct case becomes complex very quickly and the background and definitions necessary to derive all but the simplest result would unreasonably lengthen this work. Those interested in a more in depth development and discussion of the theory are encouraged to read Baumol et al.'s book, particularly Chapters 1 and 2 and Chapters 7 through 11; the intervening chapters give an in depth discussion of the determinants of market structure in multiproduct industries.

Briefly, a contestable market is a market which has the following characteristics:

i) Any potential entrant can produce a good which is a perfect substitute for the incumbent's product for the same cost.

ii) An entering firm can assume that the incumbent will not change his price in response to entry.

The last restriction can be relaxed to the point of having the incumbent's price-adjustment lag be longer than the

entry-exit lag if there are no sunk cost.

Alternatively, a contestable market must have no barriers to entry or exit. This not only means that entry and exit are legally unrestricted, but also that there are no costs involved in the process that an entrant would have to bare that an incumbent firm would not. Sunk costs can be a barrier to entry if the incumbent firm is allowed to adjust its prices, but fixed costs and scale economies are not per se entry barriers. At high levels, though, fixed costs and scale economies can guarantee sustainable natural monopolies. However, in the absence of entry barriers, these markets still cannot earn a positive profit.

If a market is perfectly contestable, then, in equilibrium, the industry output must be produced at the lowest possible cost, so that the industry must be organized efficiently and each of the firms in the industry must operate efficiently. Also, in multiproduct firms, there can be no cross subsidies amoung any subsets of the firms' product set. Firms in contestable markets can earn a profit not higher than the height of the entry barrier, which is zero in a perfectly contestable market. Lastly, in a multifirm market the price of the good must equal its marginal cost, so that the conditions for a "first best" Pareto optimum are satisfied. If, however, the industry is a natural monopoly, then the marginal cost of a good cannot equal its price so that a "second best" Pareto optimum is achieved.

In its most used form, contestable market theory predicts that all contestable industries, even natural monopolies, will exhibit competitive behavior. This specifically implies that all the firms, in multifirm markets, are producing at the minimum point on the average cost curve and that no firm is earning more than a normal profit. And even monopolists must practice average cost pricing. This statement of the results, referred to as the contestable market hypothesis, was used in the empirical and experimental work that has been done on contestable market theory (Bailey and Panzar, 1981; Coursey,Isaac, and Smith, 1984b; Coursey, Isaac, Luke, and Smith, 1984a; Harrison and McKee, 1985).

The essential contribution of contestable market theory is the idea of costless entry and exit. This makes the markets vulnerable to hit-and-run entry by firms outside the market. In contestable markets no firm can operate inefficiently or earn a postive profit, even in the presence of significant fixed costs or scale economies.

Now that the basics of contestable market theory have been presented, the remaining chapters will concentrate on how the verity of the theory might be evaluated. The next chapter will compare contestable market theory and conventional entry barrier theory. Chapter Four will present an outline of the validation process, and the penultimate chapter will present and analyze the available empirical

work. The final chapter will present some factors which must be considered in developing and applying an empirical test of the theory.

III. Contestable Markets and Entry Barrier Theory

The previous chapter developed contestable market theory, but it treated it separate from conventional entry barrier theory. This chapter will briefly present some of conventional entry barrier theory and relate this theory to contestable market theory.

Entry barrier theory, in general, consists of at least two distinct elements: assumptions about how incumbent firms will respond to entry, and a definition of an entry barrier. Based on these, the theory can distinguish what may or may not be an entry barrier and how entry barriers will effect market performance. Different assumptions can lead to very different results.

In contestable market theory an entry barrier is defined to be any cost which a potential entrant must bear that is not borne by an incumbent firm (Baumol et al., 1982a:282). This is the same definition used by Stigler (1968:67) and is equivalent to that used by Bain (1965:3). Also, contestable market theory assumes that incumbent firms will not change their prices in response to entry (Baumol et al., 1982a:5).

These assumptions require that firms in a contestable market must price at average cost (at minimum average cost in multifirm markets) in the absence of entry barriers. In the presence of entry barriers, the firms must price so as to earn a profit not higher than the value of the entry

barrier. If a firm attempted to price higher than these prices, it would be forced out of the market by a firm offering a lower price. Also, although fixed costs and scale economies can encourage sustainable configurations which are immune to entry, they do not constitute entry barriers; fixed costs must be paid by both incumbent and entering firms, and scale economies do not impose a cost differential between incumbent and entering firms if they produce at the same level of output.

The Sylos Postulate

The basis of much conventional entry barrier theory is the Sylos postulate. This postulate says that potential entrants expect incumbent firms to maintain output at pre-entry levels in the face of entry and that incumbent firms do behave this way if entry occurs (Needham, 1978:159-164). Under the Sylos postulate, if entry occurs, the incumbent firms are willing to accept the lower prices for their goods, which will result from the expansion of industry output. Also, one commonly used definition of an entry barrier in conventional entry barrier theory is that an entry barrier is any obstacle which may prevent a new firm from entering into the production of a particular good (Needham, 1978:158).

Contrast the Sylos postulate with contestable market postulate: the Sylos postulate assumes incumbent firms hold output constant, while contestable market theory postulates

that the incumbent firms will hold their price constant. As will be shown later, this will lead to a difference in the treatment of scale economies; scale economies can constitute entry barriers under the Sylos postulate.

Absolute Cost Difference Entry Barriers. The first type of entry barrier to be considered is the absolute cost difference entry barrier. This type of entry barrier arises when an entering firm's costs are higher than the costs of an incumbent firm at comparable levels of output. This type of entry barrier corresponds exactly to the definition of an entry barrier in contestable market theory. Thus, every entry barrier in a contestable market is an absolute cost difference entry barrier.

If entry barriers are present in an otherwise contestable market, the firms in the market can charge a price greater than average cost by an amount equal to the height of the entry barrier without inducing entry; that is, they can charge $\underset{\sim}{p} = E(\underset{\sim}{y}) + ATC(\underset{\sim}{y})$ (where $E(\underset{\sim}{y})$ is the height of the entry barrier) and still offer no opportunities for positive profit to entering firms. Following the same line of reasoning, under the Sylos postulate incumbent firms can charge a price equal to the average cost that an entrant would incur without inducing entry. And, since the incumbent firm's costs are less, the incumbent can earn a profit equal to the difference between his costs and an entrants costs without inducing entry.

Thus, both the Sylos postulate and contestable market

theory predict the same entry-deterring behavior in the presence of entry barriers. A difference between the postulates arises when deviations from this behavior are encountered. In contestable market theory, if a firm is earning profits greater than the height of the entry barrier, a firm could then enter the market offering a price equal to the height of the entry barrier. Since this firm can serve the same market demand as the incumbent, albeit at a higher cost, and is offering a perfect substitute at a lower price, the original firm will be forced out of the market since, by hypothesis, it cannot change its prices before the entering firm is in place in the market. Consider the same situation under the Sylos postulate. If prices are above the height of the entry barrier, an entering firm could offer goods at a price equal to the average total cost and earn a nonnegative profit. This new entry will expand output and drive prices down, but it will not change the original firm's output; the original firm and the entering firm will share (not necessarily evenly) the market at the reduced price. Thus, under contestable market theory, misbehaving firms are forced out of the market, while under the Sylos postulate they are forced only to share the market. This same difference will have a much different effect in the presence of scale economies.

Scale Economies as Entry Barriers. As has already been mentioned, scale economies do not constitute entry barriers

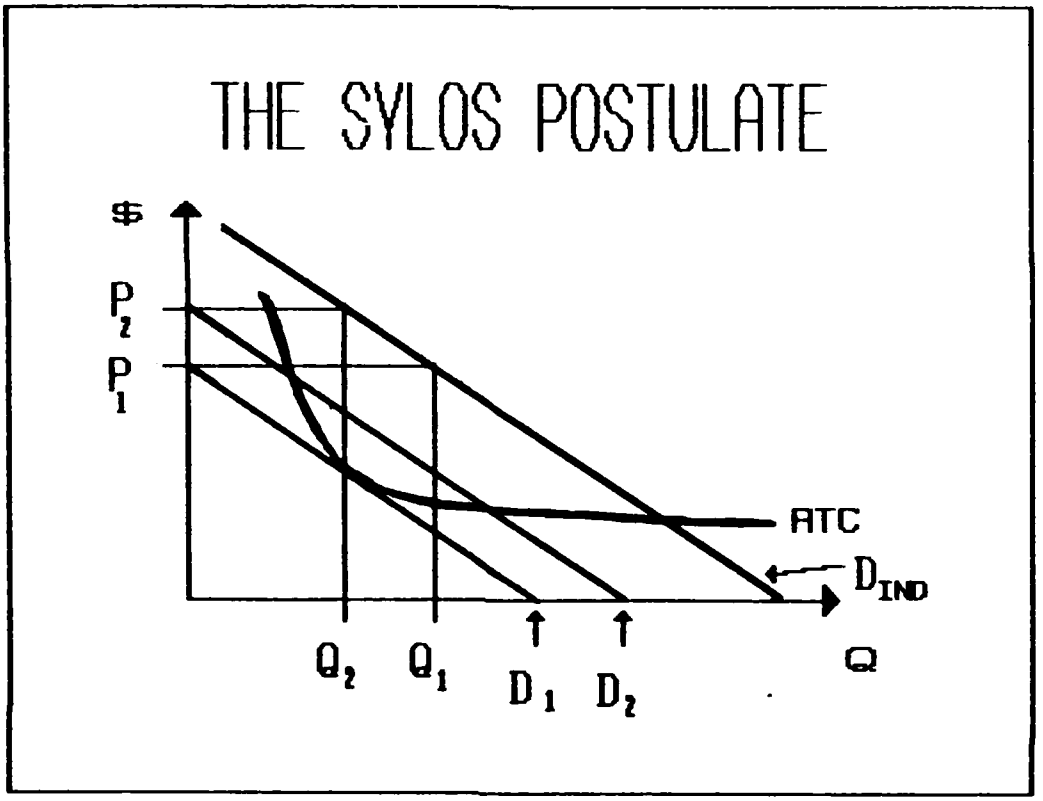


Fig 2. Scale Economies and the Sylos Postulate

(Needham, 1978:162)

in contestable markets. Under the Sylos postulate, however, economies of scale will indeed constitute entry barriers. Consider the market depicted in Figure 2. For the sake of exposition, assume that there are no absolute cost difference entry barriers in this market, so that all firms

will have the same average total cost curve, labelled ATC. As is evident there are definite scale economies. Also, let D_{IND} denote the industry demand curve.

Suppose the incumbent firm is offering output Q_1 with a price of P_1. An entering firm also offering a price of P_1 would face a demand of zero units; at P_1 the market will only support an output of Q_1 which will be supplied by the original firm, even after entry. At prices below P_1, an entering firm could supply only the difference between market demand and Q_1. Thus, if an incumbent firm is producing Q_1, an entering firm faces a demand of Q_1 units less than industry demand. D_1 represents the entrant's demand curve when the incumbent is offering Q_1, and, similarly, D_2 is the entrant's demand curve when the incumbent is offering Q_2 units of output at P_2.

Now, notice that if the incumbent firm is producing Q_2, there is a portion of the entrant's demand curve D_2 which lies above the average cost curve. Thus, an entering firm can profitably offer an output in this range and split the market with the incumbent firm; the incumbent does not deter entry at Q_2. This is not the case with demand curve D1. If the incumbent firm produces an output of Q_1 there is no portion of the entrant's demand curve which is above the average total cost curve. Thus, the incumbent firm can earn a profit equal to the difference between P_1 and $ATC(Q_1)$ without presenting an opportunity for positive profit to an entering firm. Note that no output above Q_1 presents an

opportunity for positive profit to entering firms while every output below Q_1 does. Also, as output increases at levels above Q_1 the difference between the market-clearing price and cost is diminishing, so that the maximum entry-deterring profit that the incumbent can earn is at Q_1. This maximum profit level is the height of the entry barrier, and will occur when the entrant's demand curve is tangent to the industry average total cost curve.

Under the Sylos postulate, then, scale economies will constitute entry barriers. In a contestable market, though, no such barrier will exist. If the incumbent firm does not offer the price and quantity combination where the industry demand and cost curves intersect, the firm will be forced out of the market by a firm offering this combination. The difference arises because under the Sylos postulate the incumbent firms are willing to accept lower prices in response to entry. Thus the incumbent firm can never be driven out of the market.

Summary. The difference between the Sylos postulate and contestable market theory lies in the assumed behavior of incumbent firms upon entry. The Sylos postulate assumes that the incumbent firms will maintain pre-entry output while accepting lower prices. Thus, the incumbent firm can never be driven out of the market. In contestable markets, on the other hand, the incumbent firms do not adjust their prices. Thus, if an entering firms offer a price slightly below the incumbent's, they will capture the entire market

and the incumbent will be forced out. Due to this difference, scale economies raise entry barriers under the Sylos postulate but not in contestable markets.

Condition of Entry (Bain, 1968:264-268)

In discussing the effects of entry on market performance Bain considers the condition of entry into an industry. This refers to the state of potential competition from outside sellers as evaluated by the advantages that the incumbent firms have over potential entrants as reflected by the ability of the established firms to "persistently raise their prices above a competitive level without attracting new firms to enter the industry" (Bain, 1965:3). If there are no entry barriers, then the condition of entry is free and the incumbent firms are not able to raise their prices above the competitive level. If the condition of entry is not free, then the incumbent firms are able to raise their prices above the competitive level, that is, above average cost, without attracting entry. If these firms earning a supernormal profit are in a contestable market, then there must be entry barriers to the market with a height equal to the amount of the supernormal profit. Thus, Bain's condition of entry does correspond to the definition of entry barrier used by Baumol et al. (1982a:282).

Bain's analysis of the effects of scale economies on entry (Bain, 1968:263-269) centers on the condition of the market after entry. Unlike the Sylos postulate and contest-

able market theory, Bain does not put forth one type of behavior that the incumbent firm will follow. In the presence of significant scale economies, Bian postulates three different situations:

1) An entrant enters at or above the minimum optimal scale.
2) An entrant enters with an output below the minimum optimal scale but still at a significant level, and the incumbent maintains output at the pre-entry levels.
3) An entrant enters with an ouput below the minimum optimal scale and the incumbent reduces output by an amount equal to the entrant's output, so that prices remain constant.

In Bain's (1965: 53) terminology, entry is significant if the increase in output due to that entry causes prices to drop far enough so that the effect is distinguishable from the price variations due to random variations in the market conditions. Similarly, economies of scale are significant if entry at the minimum optimal scale is significant and, if operating below the minimum optimal scale, unit costs are significantly higher than at the minimum optimal scale.

In the first case, the output of the industry will increase significantly. The incumbent firm has two options: accommodate the entrant with a reduction in output, or maintain output at or about the pre-entry level. If the incumbent firm chose to reduce output, it could reduce it

sufficiently to ensure that the price would not rise greatly, and the market demand and price would be the same as before entry. But this would involve a significant output reduction by the incumbent, which is not likely. If the incumbent does not reduce output, then the market price will fall due to the significant increase in output, and prices would sink to below the pre-entry level. Thus, if the incumbent firm set prices so that post-entry price would be low enough to guarantee zero profit to an entrant, then no entry will occur. Note that this is equivalent to the Sylos postulate; the incumbent firm is willing to accept lower prices and maintain constant output.

The second scenario, significant entry below the minimum optimal scale, leads to the same consequence. In this case, it is also assumed that the incumbent firm does not reduce output, in line with the Sylos postulate, and accepts the lower post-entry price for the good. Again, the incumbent firm can raise prices above minimum average cost. If the prices insure that the entrant firm will earn a non-positive profit at the post-entry price, no entry will occur. Since the entrant operates at a scale below the minimum optimal scale, its costs will be greater than those of the incumbent firms, so that a post-entry price higher than in the first case will suffice to insure that no opportunity for positive profit exists; the incumbent firm can elevate prices higher than in the first case. Counteracting this effect, however, is the smaller, though still significant,

price reduction due to the increase in industry output. Thus, the entry barrier may or may not be higher than in the first case, but it nonetheless exists.

In the two previous scenarios the Sylos postulate was active; that is, the incumbent firms did not decrease their outputs. In the third situation, however, the entrant accommodates the entrant by reducing output sufficiently to keep price constant. If entry is insignificant, then no price decrease due to entry will be noted, and the incumbent can maintain output at the pre-entry level. In this case, the incumbent can raise prices to the level of the entrant's average cost, which is significantly higher than the incumbent's, without attracting entry. If entry is significant but below minimum optimal scale, the incumbent can still elevate prices to the average cost of the entrant without attracting entry. If, however, the entrant enters at or above the minimum optimal scale, then the incumbent cannot raise prices above the level of minimum average cost. This corresponds to contestable market theory.

Note that in these scenarios the only variable the incumbent could control was whether or not it would reduce output. Whether or not the incumbent reduces output, the degree to which the incumbent can raise prices above the minimum average cost depended upon the level of output that it conjectures the entrant will enter with. If the incumbent chose to reduce output to keep price constant, the worst case, from the point of view of the incumbent, would be

entry at the minimum optimal scale, which yields the same results as the contestable market postulate. So, in Bain's analysis, the existence and height of the entry barriers raised by scale economies depends on the conjectures of both the incumbents and the entrants; it is not deterministic.

Summary

This chapter has reviewed conventional entry barrier theory and contrasted it with contestable market theory. The Sylos postulate assumes that the incumbent firms will maintain their output at the pre-entry level in response to entry, accepting lower prices due to the expanded output. Under the Sylos postulate scale economies will in fact raise an entry barrier. The exact height of the entry barrier will depend upon the demand conditions and upon the shape of the average cost curve. Contestable market theory, on the other hand, assumes that the incumbents will not accept lower prices for their goods. Any price above the average variable cost, in the absence of entry barriers, will cause the incumbent to be forced out of the market. Thus, scale economies will not raise entry barriers. Bain's analysis focusses on conjectures. Possible conjectures include both the Sylos and the contestable market postulates, so that the presence of entry barriers due to scale economies is not deterministic.

So, what does contestable market theory offer that is different from established theory? In markets with large

numbers of firms, contestable market theory predicts the same results as competitive theory, since competitive markets are in fact contestable. So we are left with monopolies. If there are no significant scale economies, contestable market theory yields the same results as tradtional entry barrier theory; nothing new is gained. However, if there are significant scale economies, the results do differ: contestable market theory still predicts zero profit where conventional entry barrier theory does allow a profit.

IV. Validating the Theory

Before any theory can be given wide acceptance, it must be validated. This validation is a two part process. One phase checks the internal consistency of the theory; are all the deductions logical? The other phase consists of evaluating the applicability of the theory.

In verifying the internal consistency of the theory, the primary concern is with the logical development of the arguments. In this phase, the assumptions of the theory are taken as given. Starting from these assumptions, then, a consistent argument, following the rules of logic, must proceed to develop the results. In this phase, then, the assumptions must be explicitly identified and the arguments leading to the results must be clearly traced out. This, hopefully, has been achieved in the preceding chapter.

There are other ways to verify the internal consistency of the theory in addition to tracing out the logical development. Once the assumptions have been explicitly identified, they define the operating rules of a model, which, in essence, is what a theory is. Then, all that remains to be done is to find a way to run the model and watch the development of a system within the controlled world of the model. A simulation of a model on a computer, for instance, can be very sensitive to the assumptions that govern the model, and, as such, may help in identifying hidden assumptions of the theory. On the other hand, additional assumptions might

be needed to simplify the theory so that a manageable simulation can be developed; such assumptions must be identified as such, and their effect considered in analyzing the utility of the results in verifying the theory.

Once the internal consistency of the theory has been established, the utility of the theory must be evaluated. Evaluating the utility of the theory asks if the theory can adequately describe the real world. In verifying the model (theory), Nicholson (1978:4-6) notes two general methods that can be used: the direct approach, and the indirect approach. The direct approach seeks to establish the validity of the theory by analyzing the assumptions of the theory. If the assumptions are valid and the theory is internally consistent, then the theory is valid. The indirect approach, on the other hand, concentrates more on the theory's predictions. If the theory can adequately predict real-world events, then the theory can be accepted.

The direct approach would seek real-world situations which are reasonably approximated by the assumptions of the theory. This approach, then, becomes very much concerned with the arena of applicability of the theory. There may be, however, very few, if any, real-world situations which satisfy the assumptions of the theory, after all, the theory is meant only to apprximate the reality. Consequently, analyzing the robustness of the theory is critical to this approach.

The indirect approach is concerned primarily with eval-

uating the predictive ability of the theory. This approach assumes that the theory's assumptions are valid, and then results of the theory adequately predict describe real-world situations. If the theory's predictions do in fact describe (a portion of) the real world, then the theory is assumed to be valid.

This chapter will discuss this validation process as it applies to contestable market theory. First the internal consistency of the theory will be discussed, and then both the direct and indirect approaches to the verification will be presented. When available, the experimental and empirical evidence will be briefly discussed so that its place in the overall validation process can be better appreciated. A detailed discussion of the studies will be presented is reserved for the next chapter.

Evaluating Contestable Market Theory

The principal test of the internal consistency of contestable market theory comes from a set of three laboratory experiments performed under slightly varying conditions in a natural monopoly context. With laboratory experimentation, specifically with simulation, a market can be constructed which satisfies almost exactly the assumptions of the theory. Thus, the results of the simulation should closely coincide with the predictions of the theory.

These experiments used computers to simulate the market, with human subjects taking the role of sellers and, with one

exception, buyers. The computers controlled the flow of information among the market participants (no other communication was allowed) and kept balance sheets and inventories for each of the participants. The sellers were provided with cost schedules such that any one seller's output would be capable of satisfying market demand and such that satisfying the demand with single-firm production would be cheaper than with more than one supplier; that is, the markets were natural monopolies. (Remember that this is a structural condition, not a behavioral characteristic.) Contestable market theory predicts that the firms in these markets will not earn a positive profit; the economies of scale do not form an entry barrier.

In the first of these experiments, Coursey, Isaac, and Smith (Coursey et al., 1984b) found that in the absence of entry barriers found that the contested natural monopoly markets tended to perform more like competitive markets than like monopolies; the profits were closer to zero than to the monopoly levels (Coursey et al., 1984a:111-112). An almost identical set of experiments was carried out by Harrison and McKee, except that they used a computer to simulate the buyers' behavior also. With buyer behavior also under the experimenters control, the markets can be brought even more into line with the theoretical models. Specifically, the buyers will now buy goods up to the point at which the marginal price in the current period equals the marginal utility of the good; price expectations would no longer be a

factor. (The buyers were given marginal utility schedules.) (Harrison and McKee, 1985:57). The results again show, somewhat more strongly than in CIS, that contested monopolies do tend toward competitive results (Harrison and McKee, 1985:64).

Coursey, Isaac, Luke, and Smith (CILS) repeated the CIS experiments with the addition of a sunk entry cost. In this case, the sunk cost will not pose an entry barrier because there were prices available which supported competitive quantities and still allowed a firm to earn profits sufficient to cover the the sunk cost within its lifetime (Coursey et al., 1984a:80). Their results indicate that the sunk cost did not deter entry and, consequently, the discipline of contestability still forced the market toward competitive outcomes (Coursey et al., 1984a:80).

Verifying the Model

The Direct Approach. Shortly after its advent, Shepherd (1984) questioned several aspects of the theory. Shepherd's most serious objection is to the assumption that the incumbent firms will make no price response to entry (Shepherd, 1984:576-577,580). Shepherd notes that this assumption probably is valid only for entry on such a small scale as to be ignored by the incumbent; when entry poses no threat to the incumbent. Such entry then does not force the incumbent to change his production or marketting plans, and contestable market theory loses its force. And, even if

such markets existed wherein an incumbent firm would not respond to significant entry by a firm offering prices just slightly lower than itself, Shepherd postulates that there would not be a significant number of them (Shepherd, 1984: 576-577, 578).

Shepherd also takes exception to Baumol et al.'s analysis of fixed costs. According to Shepherd, fixed costs will not impose entry barriers only when the entrant can enter on the same scale as the incumbent; high fixed costs will indeed offer a price advantage to incumbents relative to small scale producers (Shepherd, 1984:577), and, as noted above, small scale entry is the only type of entry which Shepherd believes is realistic. Shepherd also notes that markets are not as easily defined as the theory might suggest, and entry and exit barriers may not be easily distinguished (Shepherd, 1984:582).

Baumol et al. have not yet issued a reply to Shepherd's criticisms. Nonetheless, several things should be noted. Most notably is Shepherd's contention that the lack of a price response by the incumbent firms is necessary for contestability. In the discussion of time-lag contestability, the requirement was that the price-adjustment lag, must not be longer than the exit lag (given a finite entry lag). While most markets still may not fall under the dominion of time-lag contestability, they should be more plentiful than the no-response markets required by Shepherd. Also, Shepherd's identification of fixed costs being an

entry barrier is correct only if full-scale entry is not allowed. If entry size is restricted, a firm will be able to raise prices to a level equal to the average cost of production of the largest allowed entry (Shepherd, 1984: 577). If costs are decreasing, then this argument parallels the Sylos postulate, which is a different from contestable market theory.

Weitzman (1983) put forth a different criticism of the theory. He argued that, as a matter of formal theory, it is impossible to have decreasing average costs without having sunk costs; a natural monopoly cannot be perfectly contestable because there must be sunk costs (Weitzman, 1983:486). Thus contestable market theory could only be applied in markets where conventional entry barrier theory already was adequate. However we already know that contestable market theory is robust enough to handle this type of entry barrier, if the sunk costs even raise an entry barrier.

To see this, consider a simple example (Weitzman, 1983:487). Consider a single-product firm that can produce at a rate of, say, 20 units a day. Also, suppose that the firm would like to maintain an average weekly rate of production of 15 units per day, because of, say, weak demand. The firm could produce at the full rate of 20 per day for the first 15/4 days and then remain dormant for the remaining 5/4 days (See Figure 3). Thus, total production would be (15/4 days)$\cdot$20 units/day = 75 units , which is an

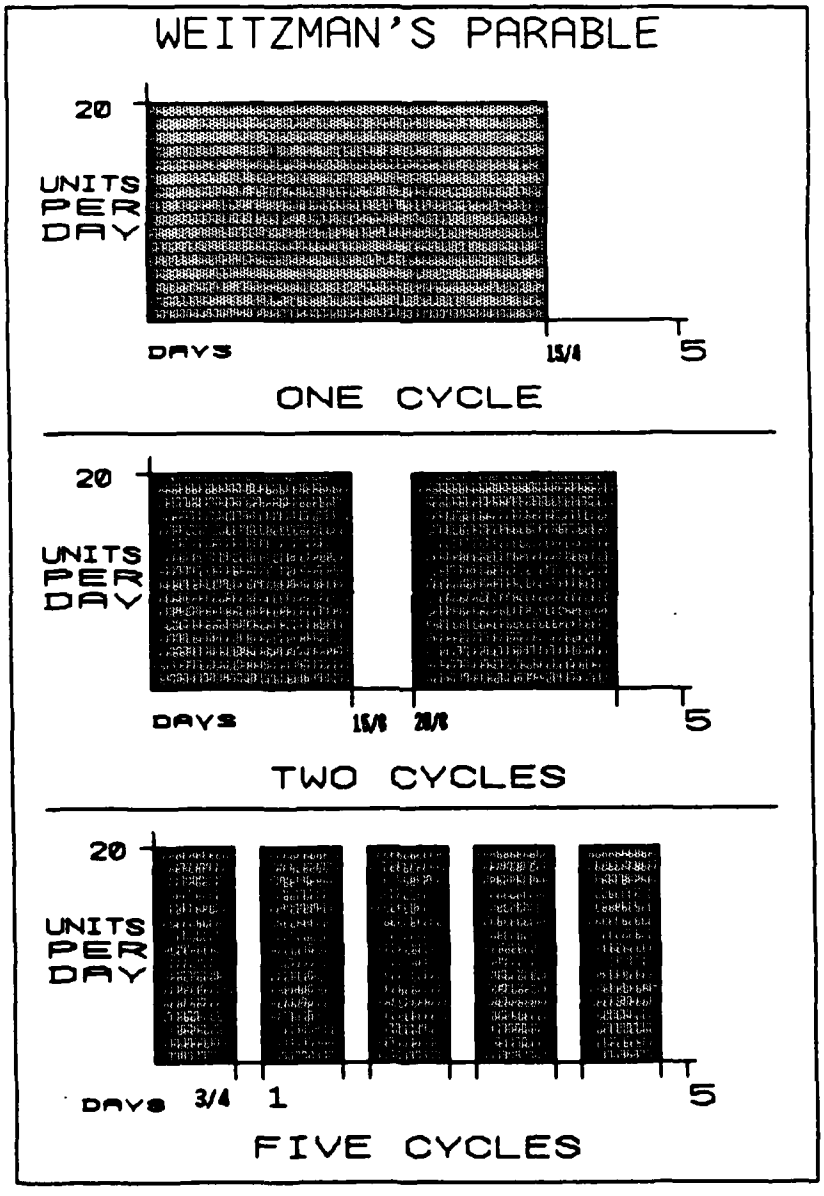


Fig 3. Possible Production Schedules

average of 15 units per day. Similarly, the firm could produce at the full rate for 15/8 days, remain dormant for a period of 5/8 days, produce at the full rate for another 15/8 days, then remain dormant for the remaining 5/8 of a day. This will still maintain the average of 15 units per day, as can be easily verified. The firm need not stop here; it can maintain the desired weekly average by continuing to decrease the length of time at which it produces at the full rate and simply increase the number of production/shutdown cycles per week. Since the firm always produces at a rate of 20 units per day when it is producing, the total cost for the weeks production is $TC = 75 \cdot C(20)$, where $C(20)$ represents the cost per unit of producing at a rate of 20 units per day. But, in the limit as the number of cycles becomes infinite, the production rate approaches a uniform rate of 15 units per day over any arbitrarily small period of time. Thus, in the limit, the total cost of production will be $TC = 75 \cdot C(15)$. If there are no costs involved in starting up or shutting down the production, that is, if there are no sunk costs, then the costs of these different production schedules must be equal. This can be true only if for any two production rates $y \leq y^1$, $C(y) = C(y^1)$, where y represents average rate of production desired to be maintained over the week, and y^1 the instantaneous rate of production. This says that in the absence of sunk costs the average cost function must be constant.

Baumol et al's (1983:493) reply picks up on several

points. First, they note that Weitzman's parable holds only for goods which can be stored at least temporarily. Thus, services must be excluded from Weitzman's conclusion. Also, goods which require an irreducible amount of time to produce (such as fermenting beer) or which require a minimum batch size for efficient production (as the length of the production/ shutdown cycles decreases, so too does the size of the lot produced at a time) cannot fit Weitzman's parable. Moreover, contestability can still maintain its discipline as long as the minimum amount of time required for efficient production, which is part of the entry lag, is less than the incumbent's price-adjustment lag, while Weitzman's argument loses its validity for any process with a nonzero minimum production time (Baumol et al., 1983:493).

Robustness. As was noted in the previous discussion of the direct approach to verification, this method must be concerned with the robustness of the theory vis-a-vis small departures from the assumptions. Contestable market theory has three principal assumptions whcih must be investigated:

1) There must be no entry barriers.
2) The price-adjustment lag must exceed the entry lag.
3) All firms, both incumbents and potential entrants, must act independently.

One of the assumptions that has received considerable attention with respect to robustness is the requirement that there be no sunk costs in a perfectly contestable market.

Shepherd remarks that virtually all production requires some type of sunk assets which cannot be sold or transferred to other uses costlessly, whether in physical equipment, advertising, *R & D*, or expert skills. Thus, hypothesizing zero sunk cost is counterfactual (Shepherd, 1984:577). But the theory is robust against departures from the zero sunk cost assumption (Baumol et al., 1983:494). First, note that from Proposition 2 sunk costs do not necessarily raise entry barriers. And, even if the sunk costs do raise entry barriers, the incumbent firms are still restricted in their pricing decisions (if they seek sustainable prices, as may reasonably be supposed); the incumbents still can earn profits not higher than the height of the entry barrier. But high sunk costs, and high fixed costs, pose an additional problem. MacDonald (1986:646) notes high fixed and sunk costs reduce the pool of potential entrants, since not all of the firms which would otherwise enter the market will be able to raise the money necessary to cover these costs. Nowhere do Baumol et al. consider this fact, and there has been no work done on the robustness of the theory vis-a-vis the number of potential entrants. As long as the firms can be kept independent, either through large numbers or by any other mechanism, the actual number of firms may not make a difference; the firms will face the Prisoners' Dilemma, and may take action which will not maximize industry profits (Nicholson, 1978:382).

Another attack on the robustness of the theory was made

by Schwartz and Reynolds (1983). Specifically, they attack the robustness of time-lag contestability, although they don't use that terminology. For instance, if the price-adjustment lag is less than the exit lag, the incumbent firm could force losses on any firm which would enter by lowering prices immediately upon entry. The incumbent firm thus can act like a monopolist. This result will hold whenever the exit lag exceeds the price-adjustment lag, even by a small amount (Schwartz and Reynolds, 1983:488-489). This is a valid criticism of time-lag contestability; it is not robust to small departures from the assumptions. Thus, contestability is not robust to departures in all of its assumptions; deviations from the noprice-response assumption can indeed destroy the power of contestability. This criticism is related to Shepherd's criticism about the reality of the nonresponse assumption.

The Indirect Approach. There are three different studies which use the indirect approach. Two of these studies (Bailey and Panzar, 1981; Allen, undated) test market performance in industries with scale economies; the third study (MacDonald, 1986) studies whether or not sunk costs do raise entry barriers. The studies by Bailey and Panzar, and MacDonald support contestable market theory; Allen's study does not support the theory.

MacDonald. MacDonald studied the effects of sunk costs on entry. Using data from 46 food related industries, MacDonald calculates the incidence of entry into and exit

from the industries over a six year period. Using entry as the dependent variable, a regression is run against several independent variables, including capital costs (which are assumed to be sunk), profitability, and growth. In addition to sunk capital costs, as measured by the value of capital in an efficiently sized plant, sunk advertising expenditures are also included. At a significance level of 95%, MacDonald reports a positive correlation of entry with growth, a negative correlation with capital costs, and no correlation with either profitability or advertising (MacDonald, 1986:646). The negative correlation between entry and sunk capital investment does indeed support the hypothesis that sunk costs can be entry barriers. The insignificant correlation between profitability and entry might seem to undermine contestable market theory, after all, profit is supposed to attract firms to the industry. First, the measure of profitability used is a measure of profit in the accounting sense, while contestable market theory always speaks in terms of normal (economic) profit. If normal profits vary amoung the industries used, then accounting profits will provide, at best, only a crude indication of which industries are earning supernormal profits, and, hence, which should be attracting entry. Further, sunk costs probably vary amoung the industries so that the amount of supernormal profit that the industry can earn without attracting entry will also vary. In short, the measure of profitability used in this is incommensurate for our pur-

poses. Also, the insignificance of advertising is counter-intuitive. MacDonald (1986:649) hypothesizes that this may be a result of grouping too many different types of industries together or to oversimplifying the role of advertising (for instance, its effect as a mobility within the industry was not examined).

Bailey and Panzar. In 1981, when the theory was still in its infancy, Elizabeth Bailey and John Panzar, who, respectively, contributed to and coauthored the 1982 book on contestable market theory, wrote a paper entitled "The Contestability of the Airline Markets During the Transition to Deregulation." In this paper they investigate the contestability of city-pair airline markets and the pricing in such markets.

First they argue that most such city-pair markets are characterized by economies of scale through the level of passengers being served (Bailey and Panzar, 1981:127). They then demonstrate that these markets can be considered contestable, although they admit that the available evidence is scanty (Bailey and Panzar, 1981:145). Thus, in the absence of entry barriers, these markets should perform competitively.

An analysis of the fare levels in these markets indicates that this is indeed the case (Bailey and Panzar, 1981:145), so that, in the absence of entry barriers and sunk costs, the evidence does support the contestable market theory. Shepherd, though, argues that not all city-pair

markets constitute valid markets. Many such city-pair are also served by alternate transportation routes that are close substitutes, so that they should also be included in the markets. Also, many of these routes play a primary role in a larger multipoint strategy, so that the decision to enter or exit a given market may be influenced by factors other than the conditions endogenous to that market (Shepherd, 1984:584).

Allen. Allen (undated) studied the profitability of shared monopolies and oligopolies. A shared monopoly is characterized by firms expecting to meet their rivals' prices and by the firms being characterized by constant returns to scale. In a shared monopoly the size of a firm is not related to its profitability. Oligopolies, on the other hand, have firms producing with either increasing or decreasing returns. In these markets, firms with similar cost structures will form groups, with each group preferring a different price.

By analyzing the price-cost margins in shared monopolies and oligopolies, Allen concludes that, amoung other things, the profitability of the firms in the oligopolies is significantly related to economies of scale. Further, the profitability of the oligopolies is higher than the profitability of shared monopolies. Allen's study, then, does not support contestable market theory. Contestable market theory predicts that profitability is not related to scale economies, which is not the case in Allen's analysis.

Summary

This chapter has reviewed the process by which contestable market theory might be validated. The internal consistency of the theory has been tested by simulation experiments. In evaluating the utility of the model, both the direct and the indiredt approaches have been used. The direct approach has been argued, with no reported experimental or empirical evidence. There is some empirical evidence availabe for the indirect approach.

All three sets of laboratory experiments to support contestable market theory. Within the framework of the experiments, scale economies did not present entry barriers. The natural monopolies did perform more like competitive markets than monopolies.

The direct approach to verifying the model has been taken by bothWeitzman and Shepherd. Weitzman argues that the theory does not apply to natural monopolies, since such markets must also have sunk costs. Shepherd argues that the assumptions of contestable market theory are too restrictive; the theory is just an analysis of a special case, not a general condition. As an additional aspect of this approach the robustness of the theory vis-a-vis small deviations from its assumptions has been analyzed. The theory is robust against the presence of entry barriers and may be robust in the face of small numbers of potential entrants. If the price-adjustment lag is shorter than the entry lag, then the theory losses all of its power.

The indirect approach receives evidence form two emprirical studies. Bailey and Panzar found that in the absence of entry barriers city-pair airline markets do perform competitively, even in the face of scale economies. Allen, however, found that in his analysis of profitability in shared monopolies and oligopolies profitability to be significantly related to scale economies. Thus, the indirect empirical evidence is conflicting. The next chapter will detail the experimental and empirical studies mentioned in this chapter.

V. Experimental and Empirical Studies

As has been said before, the evidence in support, or even against, contestable market theory is limited. The evidnece that does exists is in two forms: experimental and empirical. The experiments are actually simulations, so that the results must be interpretted in light of the assumptions made. The evidence that exists, though, does support the theory.

The Experimental Evidence

The experimental evidence in support of contestable market theory comes from three sets of laboratory experiments reported by Cousey, Isaac, and Smith (CIS), Harrison and McKee, and Cousey, Isaac, Luke, and Smith (CILS). All of these experiments tested different aspcets of the theory in a natural monopoly context. The CIS experiments tested markets with no sunk costs or entry barriers and with collusion between sellers prohibited (Coursey et al., 1984b). Then, CILS performed the same set of experiments with a sunk entry cost required from firms to enter the market (Coursey et al., 1984a). Harrison and McKee then continued with the CIS experiments, perofrming additional experiments with more potential sellers and experiments where collusion between the sellers was aloowed (Harrison and McKee, 1985).

The Posted Offer Institution (Coursey et al., 1984b: 96-99). All of the experiments used markets organized under the posted offer institution (Coursey,et al. 1984a:75-76 ; Coursey et al., 1984b:96-99; Harrison and McKee, 1985: 54-57), which will be discussed below. Some of the principal characteristics of the system include sellers posting a nonnegotiable price at which deliveries will be made in quantities demanded by the buyers, subject to seller capacity constraints; the sellers post prices independently; and the buyers themselves, rather than an auctioneer or regulatory agent, decide, through their purchases, how the sales will be distributed among the sellers. As implemented in the experiments, a computer simulated the marketplace, adjusting the balance sheets of the buyers and sellers.

In the experiments, the markets were conducted as follows. Each seller was given a marginal cost schedule, and each buyer was given a schedule of the marginal valuation of each unit purchased. These schedules need not be identical for all buyers and for all sellers, but in the experiments all of the sellers had identical marginal cost schedules. Also, the length of the experiment, given as a specifed number of trading periods, was announced to all buyers and sellers.

A trading period begins with the sellers (two sellers were used in the experiments, but the number of buyers was not specified) independently and simultaneously posting a price offer and the number of units available at that price.

No seller is allowed to post an offer which would guarantee a loss if all the offered units were sold, but every firm in the market must offer at least one unit for sale. Once all sellers have finalized their offers, the prices are shown to all participants (buyers and sellers), but the quantity offers are shown only to the sellers.

The buyers are then randomly ordered to begin their purchases. Sequentially, each buyer is allowed to make purchases up to the point at which the marginal valuation of the next unit purchased is less than its cost. The buyers may purchase any quantity from either seller, subject to the buyer and seller capacity constraints. The purchases are known only by the buyer and seller making the sales; purchases, sales, and profits are all private. After the last buyer has finished, the next trading period begins.

Each experiment operated within this general framework. The CIS experiments changed nothing. The Harrison and McKee experiments used computer-simulated buyers to allow for control of buyer demand withholding (buyers not purchasing up to the point at which marginal cost, i.e. price, equals marginal value), a problem noted in CIS (Harrison and McKee, 198557-58). They ran the same basic experiment as CIS, and performed additional experiments allowing for collusion, adding a third seller, and controlling for demand withholding. The third seller and the demand revelation (prohibition of demand withholding) were introduced in separate experiments; the permission of collusion was introduced with

three sellers. The CILS experiments added a sunk entry fee; that is, a licensing fee, good for five periods, was required to be paid to enter the market. They also conducted experiments with the buyer's demand revealed.

A necessary condition for consumers to maximize their utility is to purchase goods up to the point where the price of each good is equal to the marginal utility (valuation) of that good (Nicholson, 1978:74-76). When consumers actually purchase up to this point, demand is said to be revealed, and if they do not purchase up to this point they are withholding demand. In these experiments the effect of demand withholding is to cause decrease the power of the monopolies. With decreasing marginal costs, the last units sold are the most profitable, so that small amounts of demand withholding may drastically decrease profits (Coursey et al., 1984b:110-111). Thus, the monopolies in the experiment may not perform up to the theoretical levels if there is significant demand withholding.

Unless noted otherwise, all of the experiments were conducted under the following premises:

1. All sellers have identical marginal cost schedules (Coursey et al, 1984b:94). This corresponds to the assumption in contestable market theory that the production technology being freely available.
2. There are no barriers to entry or exit (Coursey et al., 1984b:102), except possibly in the CILS

experiment (a sunk cost need not always be an entry barrier).

3. Price responses are permitted, and are implemented immediately (that is, in the trading period for which they are calcualted). Thus, the price-adjustment lag is zero.
4. Any single seller can satisfy the market demand (Coursey et al., 1984b:94).
5. The cost schedule exhibits decreasing marginal costs to capacity(Coursey et al., 1984b:94; Harrison and McKee 55). This, combined with the ability of any single seller to satisfy market demand, implies that the markets are natural monopolies.
6. Full-scale entry is possible. Shepherd's criticism of price nonresponse in the face of significant entry does not apply to these markets since the incumbent firms are permitted to adjust prices.
7. The goods are made to order, in the sense that carrying inventory is not permitted and there are no penalties for goods offered for sale but not sold (Coursey et al., 1984b:98; Harrison and McKee, 1985:55). These markets may be viewed better as service markets. Since the goods cannot be stored, Weitzman's parable (no sunk costs imply a constant average cost) does not apply.
8. There is no explicit nonprice communication between

the buyers or the sellers (Coursey et al., 1984a: 69), except in the Harrison and McKee collusion experiments. Thus, fewer firms are needed to avoid collusion. Some examples of implicit price leadership were noted, although such actions rarely proved effective (Coursey et al., 1984b:110).

9. The buyers' marginal valuation schedules are strictly decreasing. Thus, there is a maximum market demand (which occurs at a price of zero).
10. Sellers have instant and perfect knowledge of their costs. Buyers have instant and perfect knowledge of the prices of the goods and the value of the goods to them.

Note that since all sellers in CIS and Harrison and McKee are required to offer at least one unit for sale each period, they can never actually leave the market. (Remember that natural monopoly is a structural, not behavioral, condi- tion.) However, since there is no cost for units offered for sale but not sold, all a seller would have to do to effectively leave the market is quote a price high enough to guarantee that none of the goods are sold. This type of exit (and entry by offering a reasonable price) is costless and immediate. Thus, such markets fill the necessary time-lag requirements for contestability. Also, since there are no sunk costs in either of these markets, the ability of the incumbent firm to make a price response should not reduce the contestability of the markets; the sellers still risk

nothing by entering the market.

In the CILS experiments, to participate in the market, a seller was required to purchase a market permit good for five selling periods. This entry cost is a sunk cost since, once the permit is purchased, its cost cannot be reduced even by discontinueing production altogether. The (sunk) cost of the permit was chosen to have the following properties (Coursey et al., 1984a:72):

1. If a firm achieves the monopoly price and quantity, the profit from one period will cover the sunk cost.
2. There is no price which supports the competitive quantity and recoups the sunk costs in one period.
3. There are prices which the competitive quantity such that the sunk costs can be recouped in two, three, four, or five periods.
4. There are prices which support the competitive qunatity, cover the average variable costs, but will not recover the sunk costs even in five periods.

The competitive quantity is "the largest quantity that can be sold without loss by at least one seller (that is, where average cost is less than or equal to price, or $AC(Q_c) \leq D(Q_c)$)" (Coursey et al., 1984b:99). There are several prices which can support this quantity (there are prices which could support more demand if there were more buyers in the market, so they will also support this

restricted quantity), so that there is a range of competitive prices. In the analysis, the term "the competitive price" refers to the largest price in the competitve price range (Coursey et al., 1984a:77). This definition is reasonable with demand restricted to an amount below total market seller capacity; decreasing the price cannot increase sales and will only decrease profits. With the demand and cost schedules used in the experiments, the competitive quantity is the maximum amount that can be demanded by the market.

In the CILS experiments (Coursey et al., 1984a:73), one of the two sellers was chosen as the incumbent firm. This firm was required to purchase two market permits (covering the first ten periods). After the fifth period the other firm can chose to purchase the permit and enter the market in any period. If a firm opts to leave the market, it is still permitted to observe the price offers made by the other firm. As noted in point 3 above, it is possible for the cost of the permit to be recovered over the life of the permit at competitive prices. The sunk cost, then, might not raise an entry barrier. An additional set of runs was made with demand revealed.

Results. The contestable markets hypothesis as tested in the CIS and CILS experiments, loosely stated, is that firms (sellers) in a perfectly contestable market must act as if they were competitive. In the experiments, the behavior of the sellers is measured by price and quantity. From

Proposition 7 we know that in any contestable market with two or more firms, the firms must produce at a level where average cost equals marginal price. Thus, as stated, the hypothesis is correct, as this is the same result as from a perfectly competitive market.

The experiments actually test two versions of this hypothesis: a strong version, and a weak version (Coursey et al., 1985a:77; Coursey et al., 1984b:104-105). The strong version of the hypothesis states that the price and quantity results of the contested markets, contested duopolies in CIS, will, in time, converge to the actual competitive results. The weak version of the theory simply states that while the results from the contested markets might not converge to the competitive quantities, they will be closer to the competitive results than to the monopoly results. Letting lowercase p and q denote the results from the contested markets, P_C and Q_C the theoretical competitive results, and P_M and Q_M denote the monopoly results, the hypotheses can be stated as follows:

STORNG VERSION:

$$H_0: \quad (p,q) = (P_C, Q_C) \qquad (15)$$

$$H_A: \quad (p,q) \neq (P_C, Q_C)$$

WEAK VERSION:

$$H_0: \quad p \leq (P_c + P_M)/2 \qquad (16)$$

$$q \geq (Q_c + Q_M)/2$$

$$H_A: \quad p > (P_c + P_M)/2$$

$$q < (Q_c + Q_M)/2$$

The equality in the strong version of the hypothesis should be understood to imply convergence over tiem of the observed results to the hypothesized values.

In the CIS and Harrison and McKee experiments a protected monopoly market experiment was run as a control. The CILS experiments used the CIS experiments as standards, since the same design was used in both cases. (Harrison and McKee developed an independent but equivalent computer program.) Any difference in the results between the contested and protected monopoly markets, then, would not be attributable to imperfections in the experimental design. This was also done because, as noted earlier, demand withholding may cause the even the monopolies in the experiments not to perform up to the theoretical level. Thus, the protected monopoly experiments (as opposed to a theoretical monopoly) serve as a standard against which the contested duopolies can be compared.

The CIS experiments, then, ideally should support the strong version of the contestable market hypothesis. The limited number of trading periods though may not be sufficient to observe the desired convergence, and other experimental imperfections may weaken this support. In any case,

at least the weak version of the hypothesis should receive support. In CILS, the presence of sunk costs may cause the market not to be perfectly contestable. Coursey et al. (1984a: 74-75) note several possible outcomes in such a situatuion: a monopoly solution, a solution with tacit collusion between the firms, and contestable market solutions (either strong or weak versions) with one or two firms active in the market. No one hypothesized result is singled out as being more or less likely than the others.

Harrison and Mckee use monopoly trading effectiveness to determine a markets behavior. This is defined by:

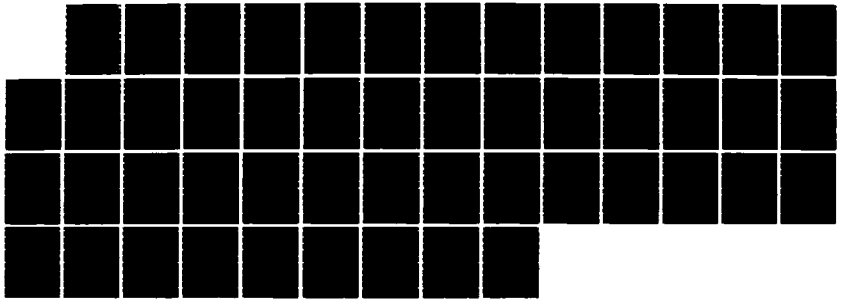
$$M = (\pi - \pi_c)/(\pi_m - \pi_c) \qquad (17)$$

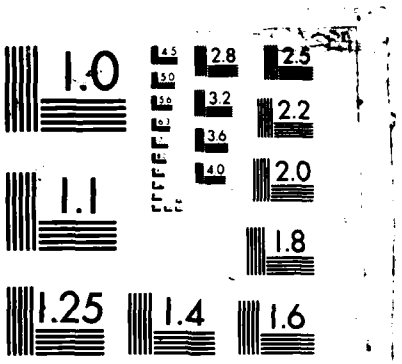
where π represents the profit of the firms in question, π_M denotes the theoretical profit of a monopolist, and π_c denotes the theoretical competitive trading profit (Harrison and Mckee, 985:59 60). CILS also report monopoly trading effectiveness, although they do not use it in any of the analysis. Sellers received a small commission for each [illegible] while this commission is included in the [illegible], even competitive markets are expected [illegible] Harrison and Mckee, 1985:55). The strong [illegible] hypothesis would then hypothesize that a [illegible] market would have a zero monopoly [illegible], and the weak version predicts that [illegible] will be less than one half (that [illegible] competitive result than to the monopoly [illegible] the number of potential sellers from

AD-A179 332 EVALUATING THE OPERATIONAL CONTENT OF CONTESTABLE MARKET THEORY(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING

2/2

UNCLASSIFIED P M LUNDY DEC 86 AFIT/GOR/OS/86D-9 F/G 5/3 NL





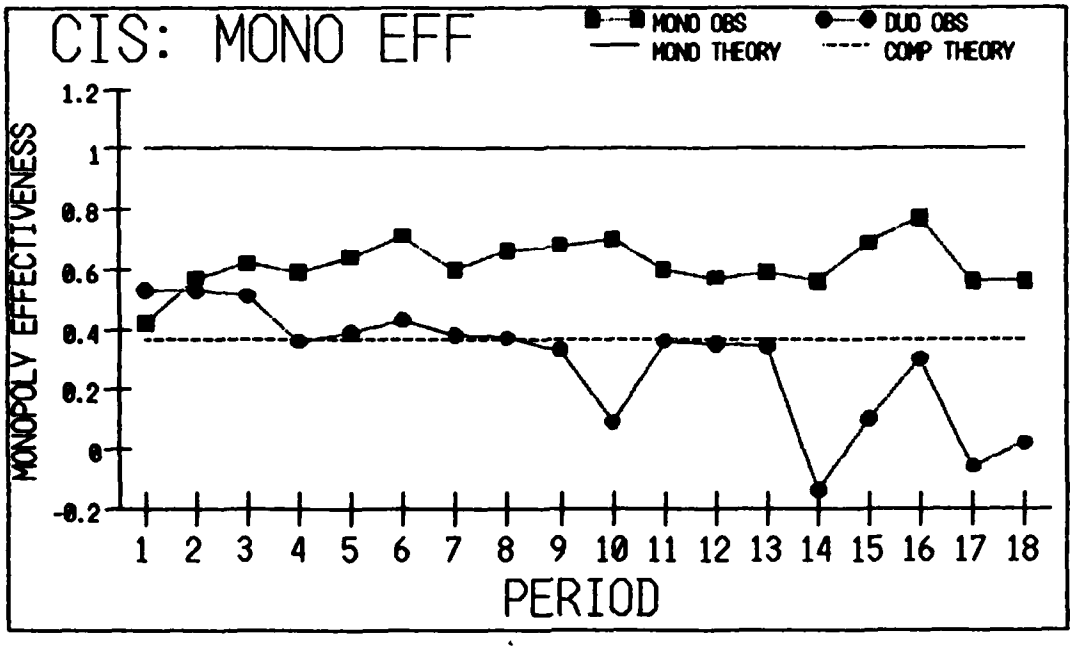


Fig 4. CIS Monopoly Trading Effectiveness

(Coursey et al., 1984b:107)

two to three should lower the market's monopoly trading effectiveness (by an undetermined amount) since implicit collusion becomes more difficult, and allowing collusion should increase the monopoly effectiveness. The exact amount by which collusion raises the monopoly trading effectiveness depends on the strength of the cartel. These experiments were run without revealing the demand.

The results of the CIS experiments are summarized in Figures 4, 5 and 6, and the Harrison and McKee results are summarized in Figures 7 and 8. The results for the

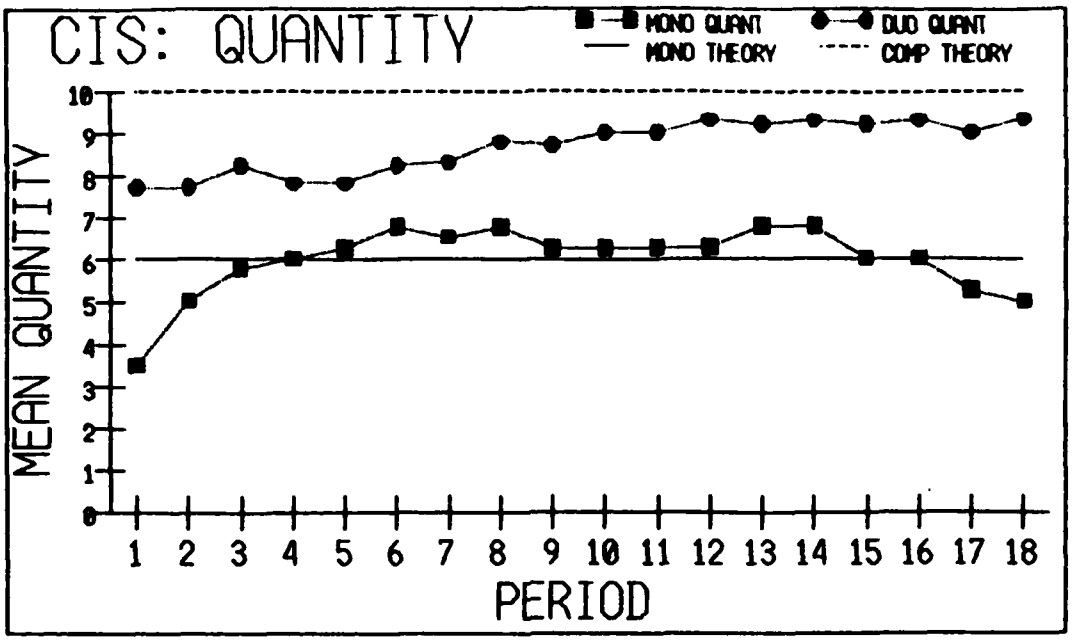


Fig 5. CIS Quantity Traded (Coursey et al., 1978b:107)

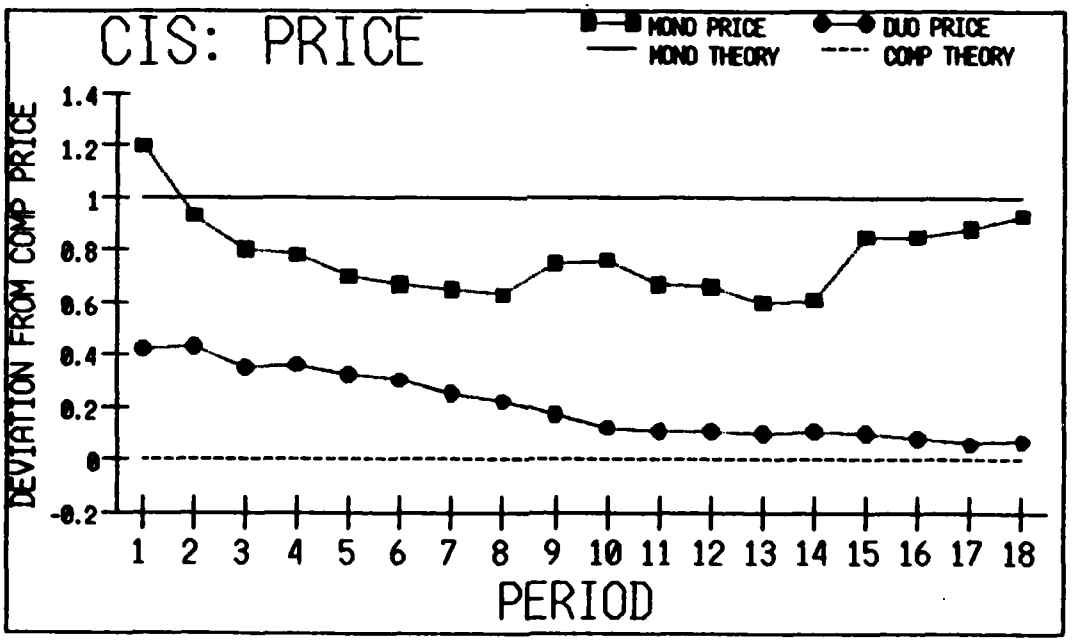


Fig 6. CIS Deviations From Competitive Price

(Coursey at al., 1984b:106)

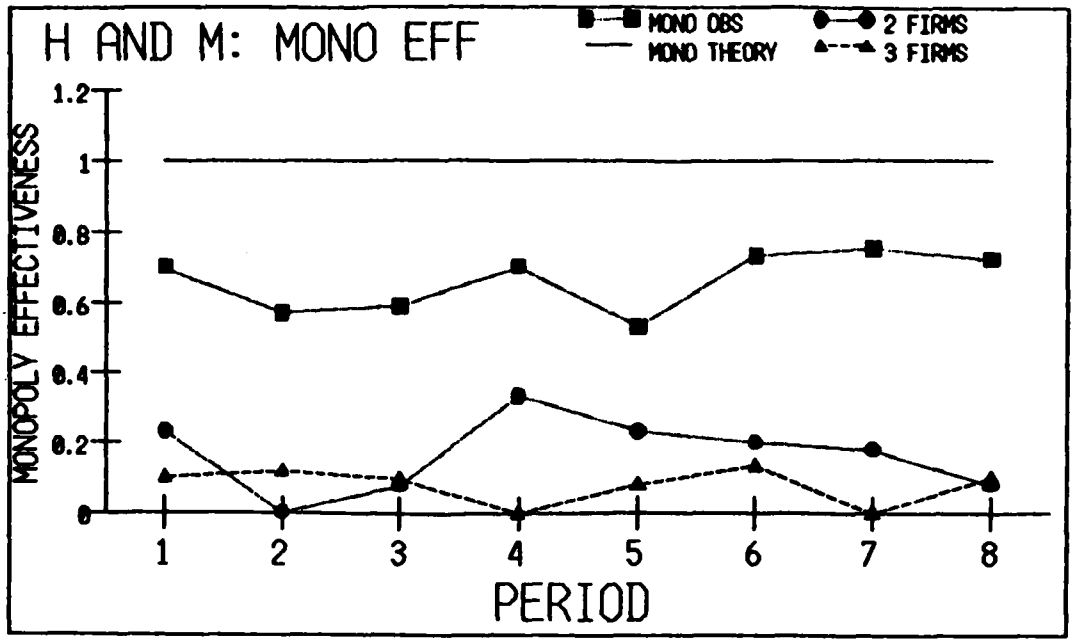


Fig 7. Harrison and McKee: 2 and 3 Firms

(Harrison and McKee, 1985:61)

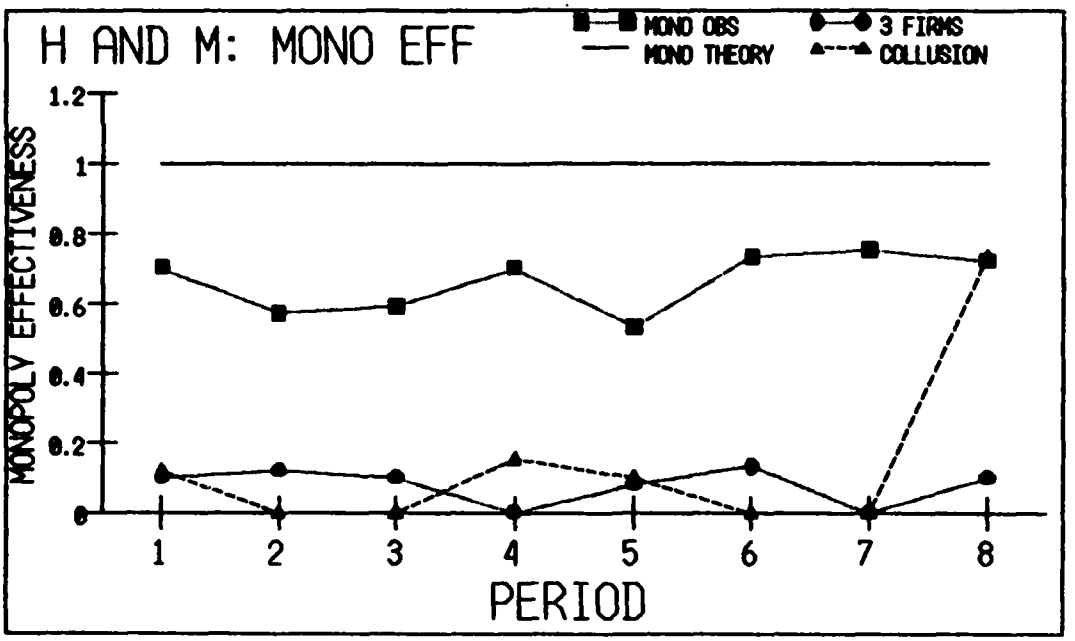


Fig 8. Harrison and McKee: 3 Firms

With and Without Collusion (Harrison and McKee 1985:61)

CIS and Harrison and McKee experiments that the contested markets always perform better than the monopoly markets, although the performance is not always up to the competitive standard.

The CIS results give clear support to the weak version of the contestable market hypothesis (Coursey et al., 1984b: 108-111). The duopoly price is always closer to the competitive price and is decaying toward that price; the duopoly quantity initially is closer to the theoretical monopoly value, but does decay toward the competitive level. The data also support the strong version of the hypothesis, but not as strongly. Of the six runs of the duopoly experiment, four converged directly to the competitive results, supporting the strong version. The other experiments were decaying to the competitive price at a rate of about 2.5 per cent per period (Coursey et al., 1984b:108-109), although no similar decay is noted in the quantity. Further, a non-parametric Mann-Whitney rank-sum test was conducted to test the hypothesis that monopoly and duopoly prices come from different distributions. This hypothesis can be accepted at the 99.99995 confidence level (Coursey et al., 1984b:110). Thus, the data clearly indicate that duopolies will perform better than monopolies.

The data also show that even the monopolies did not price at the monopoly levels. Coursey et al. (1984b: 112-113) indicate that this is due to withholding of demand, reported at 9.14 per cent in the monopoly markets. However,

since the monopolists were able to price closer to the predicted levels in the later periods, this did not appear to be a major problem. Thus, the 1.16 per cent withholding noted in the duoply experiments is not expected to have a significant effect on those results.

Figures 7 and 8 display the results for the Harrison and McKee experiments (Harrison and McKee, 1985:54). Again, the contested markets always perform better than the monopoly markets. There appear to be no significant trends in either the monopoly nor the contestable market data. Collusion appears to have little effect; at standard confidence levels, a Mann-Whitney test is unable to distinguish between the two populations. At a 92.9 per cent confidence level a Mann-Whitney test confirms that increasing the number of contestants from two to three does reduce monopoly effectiveness, and demand revelation increases the monopoly trading effectiveness of the contestable markets, as supported by a Mann-Whitney test with a 99.1 per cent confidence level (Harrison and McKee, 1985:64).

Table I summarizes the combined results of the CIS and CILS experiments. The top row labels show the different experimental designs that were run, with the no entry runs takenm from CIS. The first column lists the different hypotheses that the runs might support. The numbers in the table show the number of runs from a given experimental design which support that hypothesis. For example, when

TABLE I

Classification of Outcomes by Hypothesis and Treatment Condition
(Coursey et al., 1984a:74)

HYPOTHESIS	NO ENTRY	FREE ENTRY	K = $2	DEMAND REVEALED
MONOPOLIES				
STRONG	2	0	0	0
WEAK	4	0	0	0
TACIT COLLUSION				
MONOPOLY	NA	0	0	0
WEAK COMP	NA	0	0	1
STRONG COM	NA	0	0	2
CONTESTABLE 2 FIRMS				
STRONG	0	4	2	2
WEAK	0	6	3	2
CONTESTABLE 1 FIRM	NA	0	1	0

there was free entry, four runs supported the strong form of the contestable market hypothesis with both suppliers in the market, and six runs supported the weak version of the hypothesis with both sellers in the market.

In the table, the strong monopoly hypothesis is that the observed prices are greater than the predicted monopoly price; the weak monopoly hypothesis is that the observed prices are closer to the theoretical monopoly prices than to the theoretical competitive prices. The contestable market hypothesis with a single firm in the market applies only to the sunk cost experiment; this conditions refers to the case when only one firm remained in the market and was pricing at an entry deterring level.

The data do support the weak version of the contestable markets hypothsis (Coursey et al., 1984a:79). A binomial test of the data from the individual runs indicate that neither the weak nor the strong version of the contestable market hypothesis is supported by the protected monopoly experiments. Thus, if either of the other cases (zero or finite entry cost) support either of the hypotheses, the difference must be due to the effects of entry. Since all of the runs with either finite or zero entry costs support the weak version of the contestable markets hypothesis, entry is the disciplining factor. Based on the ruling price in the final period, all of the sunk cost experiments do support the weak version of the hypothesis; however, there is a weakening of support for the strong version of the

hypothesis (Coursey et al., 1984a:83). There was, however, a weakening of the discipline imposed by contestability in that the behavior of the market in other periods often varied greatly, instead of decreasing monotonically, as with the CIS experiments (Coursey et al., 1984a:83).

Summary. Within the context in which they were tested, the experiments with zero sunk cost all support the weak version of the contestable markets hypothesis; all of the freely contestable markets behaved more competitively than the monopolies even in the presence of scale economies. The support is not as strong for the strong version of the hypothesis. Allowing collusion did not significantly influence ability of the firms to behave as a cartel, while increasing the number of sellers in the market did have a beneficial effect. The presence of sunk costs did weaken support for the strong version of the hypothesis and, overall, weakened the competitive discipline of the markets.

These experiments were all conducted in a natural monopoly market. Full-scale entry was also permitted, so that the markets avoided Shepherd's objection. Further, the markets could be considered service oriented, and these were no entry and exit barriers. Price responses were permitted, but since the price-adjustment lag did not exceed the entry lag, the markets remained contestable.

Bailey and Panzar: The Contestability of Airline Markets

All of the evidence for contestable market theory discussed so far has been laboratory work. This type of work tests only the validity of the theory in a controlled environment. Nothing is said, though, about whether or not the assumptions are realistic or the theory can adequately predict real-world events. This is left to real world and emprirical studies.

The first empirical study of contestable market theory was conducted by Bailey and Panzar in 1981. In this study Bailey and Panzar examine the pricing in city-pair airline markets. First they argue that many such markets are natural monopolies and butress their claim with some empirical data. Next, using the prevailing route authorization policies of the Civil Aviation Board, they demonstrate that entry and exit into these markets is generally unrestricted. Then, since there are relatively little sunk costs to the airlines in such markets, these routes are contestable. The Civil Aeronautics Board does not restrict the fares of commuter airlines while it does restrict the fares of local and trunk airlines, with the restrictions on the trunks being more strict than for the locals. If these markets are contestable, then the local airlines should not being charging fares higher than the maximum rates which the trunks could charge. Using empirical data, this is shown to be the case.

The City-Pair Airline Markets (Bailey and Panzar, 1981: 125-134). A typical city-pair airline market displays economies of scales primarily because the scale economies with respect to size for individual aircraft; in a market of a given length, the cost per person can be as much as 33 per cent lower for an aircraft twice as big. Thus, there is an incentive for airlines to make fewer runs using larger aircraft in a given city-pair airline market; the equilibrium number of flights in most markets should be relatively small. Further, if the average cost per flight declines up to a few flights per day, the authors argue, most of the city-pair airline markets will be natural monopolies. This follows because the economies of scale offered by larger aircraft will encourage the airlines to use fewer flights per day, which puts the equilibrium number of flights per day on the portion of the cost curve where average cost is decreasing. To support this line of reasoning, the authors cite evidence that indicates that almost 70 per cent of the domestic nonstop markets in the United States are served by a single carrier. Thus, the hypothesis that many city-pair airline markets are natural monopolies operating in the presence of scale economies may reasonably be accepted.

Entry and exit into these markets is relatively unrestricted. Before an airline can operate over a given route, it is required to get route authority from the Civil Aeronautics Board (CAB). Starting in 1978, the CAB became less restrictive in granting authority, until, in 1979, all

requested route authority was granted, unless there were environmental or noise restrictions at the airports or the carrier's fitness was in question. Once an airline had authority to operate in a given route, it was not required to actually operate that route; they could enter and exit the authorized markets as they pleased.

Also, there are few sunk costs to an airline entering a given city-pair market. An airplane can be quickly moved from one city-pair route to another, so it is not sunk capital. The capital costs associated with airports are certainly sunk, as they cannot be recovered except in the long-run. These costs, however, are not borne by the airlines, but by the municipalities they serve. Thus, since there are few sunk costs associated with a given city-pair market, and entry and exit from the markets is generally unrestricted, these markets should be contestable.

The above argument has ignored several market imperfections. The environmental and noise restrictions enforced by some airports have already been mentioned. Overcrowding at some airports may also restrict entry. Also, in an effort to underwrite the sunk costs of the airports, some municipalities bind certain airlines under long-term leases for gate and terminal space. In an extreme case, this could give an airline a degree of monopoly power over the airport, as it could decide when, to whom, and at what price it will sublease. The authors contend, though, that all these market imperfections do not seem to have had much effect on

entry. In the year starting 2 July 1978, there were 100 cases of new entry into 143 hub airports (Bailey and Panzar, 1981:134).

Results (Bailey and Panzar, 1981:141-145). So far we have established that most of the city-pair airline markets at least approximate contestable markets. Sunk costs are minimal, and, generally, entry into and exit from these markets is relatively unrestricted. Also, most of these markets are natural monopolies. Since they are contestable, though, they should approximate competitive performance. The measure of performance used in this study is the fare level.

Before the design of the study can be presented, the CAB's pricing policy must be briefly explained. The explanation given here will briefly summarize that given in Bailey and Panzar (1981:134-137), which is itself drawn from various CAB documents. Most of the calculations of fare levels are based on the Standard Industry Fare Level (SIFL), which is approximately equal to the standard coach fare in 1977. This fare is given in terms of fare per mile, and is itself a function of the route length. Starting in mid-1977, the CAB began giving carriers a degree of downward pricing flexibility, allowing, in general, reductions of up to 50 per cent without justification.

Upward pricing flexibility, though, is more limited. The degree of flexibility permitted in a market to a carrier depends on the level of competition within the market, the

length of the route, and whether the carrier was a trunk airline or a local carrier. Trunk airlines were allowed to raise their fares 10 per cent above the SIFL fares in markets with four or more authorized carriers, but only 5 per cent in markets with fewer authorized carriers. An additional 5 per cent increase is allowed to nonmonopoly carriers (carriers which transport less than 70 per cent of the passengers between two points). The formula used to calculate the SIFL is known to discriminate against short-haul (less than 400 miles) market. So to alleviate this the CAB allowed local carriers, whose routes averaged about 200 miles, to price at 130 per cent of the SIFL, regardless of the length of the route.

At the time the study was done, the supply of aircraft was having a definite effect on the pricing strategy. Production lags limited the available aircraft to less than the number demanded, so that prices would not be driven to the lower bounds permitted by regulation. But, the supply was also great enough to prevent the upper fare limits from being reached. Trunk carriers could, on the average, expect to earn approximately 105 per cent of the SIFL on any given route. That is, if a market could not support a price of 105 per cent of the SIFL, then a trunk carrier could transfer the aircraft to another market and expect to earn that much. Thus, if local carriers charged more than 105 per cent of the SIFL on a route for which a trunk carrier could compete (some small markets might not be able to

support the trunk's larger planes), the trunk would find it profitable to transfer planes to that route.

The research strategy, then, is to find markets for which both local and trunk carriers would compete. The local carriers would be permitted to price these routes at 130 percent of the SIFL while the trunks could only price at about 110 per cent of the SIFL. However, the local carriers should not price over the 105 per cent of the SIFL due to the contestability of the markets.

Table 2 shows the results of a regression by the CAB's Office of Economic Analysis (Bailey and Panzar, 1981:142) to test this hypothesis. The sample consisted of 43 markets, ranging in length from 85 to 778 miles, served by local carriers. Both monopoly and competitive markets were included. The dependent variable is the local carrier's fare (if there were multiple carriers, the carrier with the most passengers) divided by the SIFL fare. The independent variables are the length of route and the presence of competition. The route lengths are grouped into blocks of 100 miles with indicator variables being used (1, if the route length is in the specified block; 0 otherwise). Three types of competition are considered: competition by a trunk in markets not greater than 400 miles in length, competition by a trunk carrier in a market over 400 miles in length, and competition by other local and commuter carriers. (The 400 mile point is used as a dividing line since the fare

TABLE II: (Bailey and Panzar, 1981:142)

Regression Results: Fares versus Route Length and Competition

Dependent Variable: Published Fare / SIFL		
Independent Variable	Coefficient	T Statistic
Constant	0.988	14.4
Length < 100 miles	0.224	2.6
101 - 200 miles	0.198	2.6
201 - 300 miles	0.195	2.5
301 - 400 miles	0.215	2.6
401 - 500 miles	0.057	0.8
501 - 600 miles	0.072	0.7
601 - 700 miles	0.023	0.3
Trunk Comp ≤ 400 miles	-0.157	-4.1
Trunk comp > 400 miles	-0.012	-0.2
Other Competition	0.03	0.9
R^2 = 0.432	Standard Error of Estimate = 0.088	

structure mandated for trunks significantly underprices such markets.) Again, indicator variables are used, with a 1 indicating the presence of such competition.

The first thing to note about the results is that none of the coefficients of the mileage variables are significant (at standard confidence levels) for ranges over 400 miles, while they are all significant at ranges below this level. Additionally, note that in the markets over 400 miles in

length, the effect of actual trunk competition is insignificant. Since the fare levels in these markets are not much greater than 105 per cent of the SIFL, potential competition, then, must be disciplining these markets. In the shorter markets, the actual presence of trunk carriers has a significant negative effect on the price, bringing the fares down to approximately 105 percent of the SIFL (Bailey and Panzar, 1981:142-143). The presence of competition by commuter and other local airlines has an insignificant effect on the fare level.

The data indicate that markets of less than 400 miles in length are distinctly different from longer markets. Part of reason for this lies in the discriminatory fare levels in these markets, with the fare limits for local carriers significantly higher than that for trunks. Another factor differentiating these markets is the special equipment characteristizing carriers in these markets. The authors note that the "specialized equipment requirements [for these markets] do not make [them] as vulnerable to entry by trunks" (Bailey and Panzar, 1981:143), although they do not elaborate. In these markets, then, the pool of potential entrants is considerably reduced, allowing higher fares. Commuter carriers did not prove to be adequate competition, perhaps because of the perceived quality differences in the services (jet versus commuter aricraft) (Bailey and Panzar, 1981:143).

In markets over 400 miles in length, the aircraft used

by trunk carriers are suitable competition for the local carriers. Thus, the local carriers were not able to raise their fares above the level which these aricraft could earn in other markets. This is what is expected in a set of contestable markets with a limited supply.

Criticism. As alluded to earlier, this study has been criticized by Shepherd (1984). His main criticism concerns the definition of the markets. He argues that most route changes are not in response to profits being earned in a particular city-pair market, but are rather part of a much larger competitive strategy (Shepherd, 1984:584). In essence, the appropriate market would then be much larger than just one city-pair. Although the decision to enter or exit a particular city-pair market is probably more complex than is assumed by Bailey and Panzar (that is, based soley on the profits available in those markets), profitability in such markets must certainly play some role in the competitive strategy that Shepherd suggests (at least some of the carriers routes must be profitable). Further, if most of the city-pair markets are contestable, then the discussion of cross subsidies comes into play: If all the markets within which a multiproduct firm operates are contestable, then the firm must avoid cross subsidies between any subset of its products. In the context of the airlines, a multiproduct firm would be a carrier authorized to operate on more than one city-pair route. It should also be remembered that this result is not very robust from departures; each market with-

in which the firm operates that is protected by an entry barrier can provide a corss subsidy to any other (possibly contestable) market, allowing the firm to sell goods in that market for less than cost.

Even if the mutlistage aspect of the market is disregarded, Shepherd remarks that there are close substitutes available to many of the city-pair markets (Shepherd, 1984: 584). The immediate response to this remark is to question the nature of those substitutes. The discipline of potential competition proved to be effective only in markets over 400 miles in length. Thus, travel by neither train nor interstate highway would seem to be a close substitute. Cities which are close together, such as Newark and New York, might provide close substitutes if both cities are the points of departure or both are points of arrival. No indication is given as to the location or the dispersion of the markets used in the Bailey and Panzar study, so nothing definite can be said in this respect. Nonetheless, it does not seem likely that a significant number of the cities would be close enough to qualify as substitutes. If this were the case however, a new study could easily resolve that issue.

There are other weaknesses in the study though. Most notable is the fact that entry into a market is not as instantaneous as simply rerouting an airplane (Shepherd, 1984: 584). Establishing ground facilities and building up patronage takes time, almost surely less time than a carrier needs

to affect a price change. Thus, contestability may not be the disciplining force. Further, the main influences on market shares stem from the interactions amoung the firms already in the market, rather from entering and exiting firms (Shepherd, 1984:585).

Summary. Bailey and Panzar's analysis of city-pair airline markets does support contestable market theory. These markets are natural monopolies with economies of scale. These markets also have small sunk costs and only minor entry and exit restrictions. The restrictions have not greatly impeded entry and exit into these markets, so that these markets are do approximate contestable markets.

Even in the presence of scale economies, the fares in these markets are approximately at the competitive levels, as the theory would predict; standard entry barrier theory would permit the local carriers to earn supernormal profits. There may be problems with defining the appropriate markets, and there may be factors involved in a carriers decision to enter or exit these markets other than just the fare level in that particular market.

Oligopolies and Shared Monopolies

The last bit of empirical evidence comes from an indirect source. Allen (undated) studied the relation between shared monopoly, oligopoly, and price-cost margins (these terms will be defined later), concluding that the two industry structures are distinct. In the comments on the paper, the

referee suggests that the structures used may present a sample which at least approximate conditions under which contestable market theory may be tested (Allen, undated b).

Allen defines a shared monopoly to be an industry with the following characteristics:

1) Every seller expects to meet its rivals' price so that any sellers demand is $q^i(p) = a^i Q(p)$, where a^i is that firm i's market share, and either
2) the firms are producing and selling output with constant returns to firm size, or
3) if firms are producing with either increasing decreasing returns to firm size, then all the firms have equal market shares.

Under these conditions, the industry's price-output solution will be a monopoly solution. In the constant cost case, the case which is analyzed, this result will hold regardless of the number or size distribution of firms in the industry as long as the actions of any one firm will influence the price. Every firm prefers the same price (since their costs are identical) which is the monopoly price, and all firms share equal market power.

Oligopolies, as defined by Allen, are characterized by firms of differing sizes producing under either increasing or decreasing cost conditions. No single price will be preferred by all of the firms; for example, with decreasing costs the firms with larger outputs, and hence smaller costs, will prefer prices lower than those preferred by the

smaller firms. This can lead to firms forming groups according to their sizes, with each group preferring a different price. Scale economies as well as other mobility barriers will help keep additional firms from entering these groups.

As noted, each group will prefer a different price, based on its costs. The influence that a particular group will have in determing the industry price will depend on the dominance of that group within the market; the more dominant the group, the greater its price influence. Thus, the more dominant groups are more likely to be profitable, since they ae able to exert greater influence on the price. Also, the more dominant groups are likely to be composed of firms with larger market shares, and hence lower costs, further increasing the likelihood of profitability.

Shared Monopolies, Oligopolies, and Contestability. The most important difference between shared monopolies and oligopolies is the presence of groups in the oligopies. These groups, based on long run cost, prefer different prices due to scale economies. The more dominant groups will probably earn postive profits, because of these economies, but they will not be vulnerable to entry because of mobility barriers. Thus, unlike shared monopolies where no such group structure exists, the profitability of an oligopoly will be dependent on the size distribution of firms within the industry.

As the referee's comments suggest (Allen, undated b),

these market structures do provide an arena within which contestable market theory might be tested. As often stated before, scale economies are not entry barriers in contestable markets. Thus, if scale economies are indeed the primary source of mobility barriers, then contestable market theory would predict that the firms in the oligopolies would not show any greater profitability than the firms in the shared monopolies.

To see this more clearly, first consider the oligopoly. Consider each group within the market to be a specific sub-industry serving a submarket. Then all firms, inside and outside market, that are not in a specific submarket are potential entrants into that submarket. In an oligopoly, these submarkets are differentiated by scale economies. Thus, if scale economies do not form entry barriers, then the group structure should not effect the profitabilty of the individual groups. Profitability thus will be independent of group structure and all firms will be equally profitable. Without the mobility barriers provided by scale economies, the oligopoly is identical to the shared monopoly.

Results (Allen, undated:6-16). The first step in this study is to find a method to distinguish which industries are shared monopolies and which are oligopolies. These two structures differ in terms of long-run average costs: shared monopolies are constant cost industries and oligopolies are either increasing or decreasing cost industries. The

measure that Allen uses is the cost advantage ratio. This is a measure of the cost advantage that the four largest firms in an industry have relative to the next four largest firms. On this basis, Allen distinguishes 15 per cent as the critical value; that is, if difference in unit costs between the first four and second four largest firms the industry is considered an oligopoly, otherwise the industry is considered a shared monopoly.

Once the industries have been classified as either shared monopolies or oligopolies, the factors which effect profitability must be analyzed. Allen considers two measures of group dominance, a measure of scale economies, the capital-output ratio, an advertising variable, industry growth and geographic dispersion. In Allen's analysis, the measures of group dominance and scale economies should be significantly related to industry profitability in the oligopolies. Contestable market theory, though, expects no significant relation between scale economies or group dominance and profitability. Also, if capital costs, as represented by the capital-output ratio, are assumed to be sunk, as in MacDonald (1986), then contestable market theory would predict that profitability should be positively related to the capital-output ratio.

After running a linear regression on profitability as a function of the indicated independent variables, Allen reports the results summarized in Table III. In the

TABLE III

Regression Results: Profitability and Structural Influences

(Allen, undated:16)

Independent Variable: Profit Margin		
Structural Influence	Relation to Profitability in:	
	Oligopoly	Shared Mono.
Group Dominance	S(1) P	INSIGNIFICANT
Scale Economies	S(10) P	INSIGNIFICANT
Capital-Output Ratio	S(10) P	S(10) P
Advertising	S(10) P	INSIGNIFICANT
Growth	INSIGNIFICANT	INSIGNIFICANT
Geographic Dispersion	S(5) N	INSIGNIFICANT
S(i) = Significant Correlation at i per cent		
P = Positive Correlation N = Negative Correlation		

oligopoly markets, both group dominance and scale economies are significantly related to profitability, which contradicts contestable market theory. Thus, contestable market theory is not a good model for the markets studied by Allen. Nonetheless, note that the capital-output ratio, which may represent sunk costs, is significantly related to profitability in both the shared monopoly and oligopoly markets. This is in accord with MacDonald's (1986) findings that sunk costs can indeed raise entry barriers.

Summary

This chapter has reviewed the experimental and empirical evidence on contestable market theory. The experimental evidence comes from simulations run with contestable decreasing-cost markets. Under these conditions, did perform closer to the competitive levels than to the monopoly levels. Thus the experiments do support contestable market theory, since conventional entry barrier theory would predict that the firms could earn monopoly profits.

The empirical evidence, which can be used to indirectly verify contestable market theory, comes form two studies. Bailey and Panzar (1981) find that local carriers serving city-pair airline markets do price competitively, even though the markets are natural monopolies. Several criticisms have been levelled against this study however. Another study by Allen (undated) analyzed profitability in shared monopoly and oligopoly markets. The essential difference between these markets is the presence of scale economies in the oligopoly markets. Profitability in the oligopoly markets was found to be significantly correlated to scale economies, which contradicts contestable market theory. Thus, contestable market theory is not an appropriate model with which to analyze these markets.

VI. Conclusion: Developing an Empirical Test

The previous chapters have discussed contestable market theory and contrasted it with conventional entry barrier theory, discussed how the theory might be validated, and reviewed existing evidence supporting the theory. Now it is time to put together what has been learned form the previous chapters: How can contestable market theory best be tested?

The Direct Approach: Verifying the Assumptions

As noted before, the direct approach to verifying a model seeks to establish the verity of the model's assumptions; it seeks real world situations which adequately fit the theory's assumptions. This approach will be difficult, if not impossible, with contestable market theory. Consider the three requirements of a perfectly contestable market:

i) Free entry and exit,

ii) The price-adjustment lag must exceed the entry lag, and

iii) There must be a large number of potential entrants.

Freedom of entry and exit is not too troublesome a condition. Neither fixed costs nor scale economies will be entry barriers, so they need not be considered (if the fixed costs are not too high). Sunk costs will pose an entry barrier, since there is usually no guarantee that the firm will be in the market for the lifetime of the investment. Even if there are entry barriers, though, the theory is

robust enough in this assumption to handle this imperfection. High fixed or sunk costs may raise an entry barrier if firms cannot obtain the necessary money to cover these costs at a reasonable price.

The large number requirement for potential entrants should not pose much of a problem either. The necessary condition in this assumption is that the firms, both incumbents and entrants, act independently. Thus, there may be a fair number of markets which fill this condition.

The price-adjustment assumption will be more difficult to deal with. The assumption will probably be satisfied in only a limited number of markets, since it is usually easier and quicker for firms to adjust prices than to enter a market. And to further complicate the issue, contestability is not robust in this assumption. Schwartz and Reynolds (1983:488-489) have shown that contestability completely loses its power if the price-adjustment lag does not exceed the entry lag.

The most common situation where price adjustment may not occur in response to entry is in growing industries. In this situation, the growth in demand is sufficient to absorb to increase in supply so that no price adjustment is needed by the incumbent firm. Thus, industries undergoing expansion may provide an arena within which contestable market theory is applicable. But in this case, contestable market theory adds nothing new to conventional theory.

The above conditions will guarantee the contestability

of the market if all of the firms in the market are single-product firms. If any firm in a market is a multiproduct firm, then either every market within which that firm operates must be contestable or the firm must be prohibited from cross subsidizing. The ripple effects of an imperfection in one market can damage the contestability of all the markets within which the firm operates. Thus, the difficulties of finding a single contestable market are exponentially complicated in multiproduct markets.

So, there does not appear to be much hope of directly verifiying contestable market theory. Even within the bounds which robustness can handle, the requirements of contestability are too stringent. Thus, we must try a different approach.

The Indirect Approach: Verifying the Predictions

In using the indirect approach, we seek situations where contestability predicts results different from conventional analysis. Multifirm markets (that is, markets which are structurely are not natural monopolies) will not be useful in this respect; contestability predicts the same results as standard entry barrier theory. Thus, natural monopolies must be the place to look. But not even all natural monopolies will work. If there are no significant economies of scale then contestable market theory predicts the same outcome as conventional entry barrier theory. So, the structure that can best be used to test contestable

market theory are natural monopolies with significant scale economies.

Within this context, what can be tested? In the presence of significant scale economies conventional entry barrier theory predicts that the incumbent firms can earn a positive profit without attracting entry. Contestable market theory predicts the opposite: any supernormal profits will attract entry. So this could be one possible test.

Upon further consideration, though, there may be other structures within which contestable market theory may be tested. Allen's study (undated) provides one such example, in which a group of firms is analog to the single firms in the natural monopoly markets. Other similar examples may exist.

All the above discussion has centered on the fact that no firm in a contestable market can earn a positvie profit. An important corolary to this proposition is that, in a contestable market, profit is independent of firm size. Numerous studies have shown that market share and profitability are strongly related, contradicting contestable market theory.

Summary

Thus, contestable market theory faces many difficulties in being verified. Direct verification appears to be a fruitless task due to excessively stringent assumptions. Indirect verification is a better approach. A natural

monopoly with scale economies would be the ideal market structure within which to test the theory, since in most other cases the theory yields the same results as conventional theory. The evidence that does exist, though, is not promising. Bailey and Panzar's study (1981) supports the theory while Allen's analysis (undated) does not. Also the well documented relation between market share and profitability do not support the theory. Thus, it appears both from the discussion of direct verification and from the few studies that have been done that the theory is not general enough. Contestable market theory is not useful as a general tool for analyzing real-world markets, although it may be useful in some special cases

Original Contributions

The following original contributions come from this research:

1) Time-Lag Contestability. In Baumol et al.'s book (1982a) and in the rest of the published literature, contestable markets are required to prohibit incumbent firms from issueing any price response. Contestability has been shown to retain its power, though, as long as the price-adjustment lag is longer than the entrance lag. This extension has implications in interpretting the criticism by Shepherd (1984) and Weitzman (1983), and in interpretting the results from the experiments

(Coursey et al., 1984a; Coursey et al., 1984b; Harrison and McKee, 1985).

2) Cross Subsidies and Sontestable Markets. The discussion of the propagation of market imperfections to any market within which a multiproduct firm operates has not appeared before.

3) Interpretation of the Experimental Evidence. Coursey et al.'s (1984a; 1984b) and Harrison and McKee's (1985) explanations are unclear and incomplete. They do not explain the mechanism by which firms enter and leave the market is lacking. Further, Harrison and McKee (1985:66) remark that the experiments neglect the price nonresponse assumption, and note that operationalizing this assumption can only strenghten their results. But, with the extension of the price nonresponse assumption to the price-adjustment lag assumption, the prohibition of pricing responses should not change the predicted results, although including such a prohibition may influence the rate of convergence of the experimental results to the predicted results.

3) Analysis of Shepherd's Criticism. Baumol et al. have not yet issued a response to Shepherd's (1984) criticisms.

4) Interpretation of Allen's Studies. The referee's comments (Allen, undated b) suggest that Allen's

analysis of oligopolies and shared monopolies may present an arena within which to test contestable market theory. The referee notes that the results may be interpretted as a verification of contestable market theory. The analysis in this research suggest the exact opposite.

APPENDICES

Appendix A: The Propositions of Contestability

The following propositions were proven in Chapter Two:

PROPOSITION 1: (Baumol et al., 1982a:289-290)

Fixed costs are not entry barriers.

PROPOSITION 2: (Baumol et al., 1982a:290-291)

The presence of sunk costs can be a barrier to entry.

PROPOSITION 3:

In a contestable market it is necessary that the entry lag be finite, and either:

1) the price-adjustment lag is greater than the entry lag if the exit lag is finite, or
2) the price-adjustment lag is infinite if the exit lag is infinite.

PROPOSITION 4: (Baumol et al., 1983:495)

Any industry configuration in equilibrium must be sustainable, but not every sustainable configuration need be in equilibrium.

PROPOSITION 5: (Baumol et al., 1982a:314-316)

In a contestable market, any equilibrium configuration must be efficient.

PROPOSITION 6: (Baumol et al., 1982a:314)

Any firm in a sustainable configuration in a perfectly contestable market must have the following properties:

(i) The firm must operate efficiently; it must produce its output at minimum cost.

(ii) The firm must earn zero economic profit.

(iii) The firm must select an output vector such that if all outputs are reduced by a factor of k, total costs are reduced by a factor of not more than k ($0 < k < 1$).

(iv) In a multiproduct market, a firm must avoid cross subsidies among any subset of its product set.

COROLLARY 6(ii)A: (Baumol et al.,1982a:201)

In a sustainable configuration in a contestable market with entry barriers, incumbent firms cannot earn profits greater than the value of the entry barrier.

PROPOSITION 7: (Baumol et al., 1982a:317)

In any contestable market with two or more firms, each firm in a sustainable configuration must select an output at which the marginal costs of production are equal to the price of the good.

Appendix B: The Existence of Sustainable Configurations

Chapter Two developed, amoung other things, some of the properties of sustainable configurations. The question of the existence of sustainable solutions was never discussed; it was assumed that a sustainable configuration existed for the market under consideration. This appendix will discuss some issues concerning the existence of sustainable configurations.

To recapitulate some of the properties developed in Chapter Two, it was shown that any sustainable configuration must minimize the total cost of producing the industry output and that any firm in a sustainable configuration must operate efficiently and earn zero economic profit. In multifirm configurations it is also necessary that price equal marginal cost, so we have $p = MC = AC$. By the definition of sustainability, any deviation from these conditions results in the opportunity for a positive profit.

This appendix will concentrate primarily on conditions sufficient to insure the existence of sustainable configurations in single-output markets. The complications introduced in multiproduct markets are considerable and, for for the most part, beyond the scope of this effort. Two different market conditions will be discussed: U-shaped short-run average cost curves with multifirm markets, and flat-bottomed short-run average cost curves in multifirm markets.

Case 1: U-Shaped Average Cost Curves in Multifirm Markets

The starting point for the discussion will be in a market within which the firms have U-shaped short-run average cost curves and demand sufficient to support more than one firm, specifically, the market is not a natural monopoly. A U-shaped short-run average cost curve is strictly decreasing as output increases up to a unique cost-minimizing output, and then strictly increases as output increases further. The important point is that there is only one output level which minimizes short-run average cost, which will be denoted by y_M. Thus, $AC(y_M) < AC(y)$ for any $y = y_M$.

U-shaped short-run average cost curves are frequently studied in microeconomic theory (Nicholson 1978:231-238). These curves arise when diminshing marginal productivity vis-a-vis inputs are assumed to set in beyond a certain point. Thus, at high input levels increasing output by one unit requires relatively more input than is required at moderate levels. At constant input prices, then, the average cost will begin to rise.

Assume that all the firms (both incumbent firms and potential entrants) in a contestable market will use the same production technology and face the same input costs. This option is permitted because the technology is assumed to be freely available. The short run average cost curve will have the shape indicated in Figure 9.

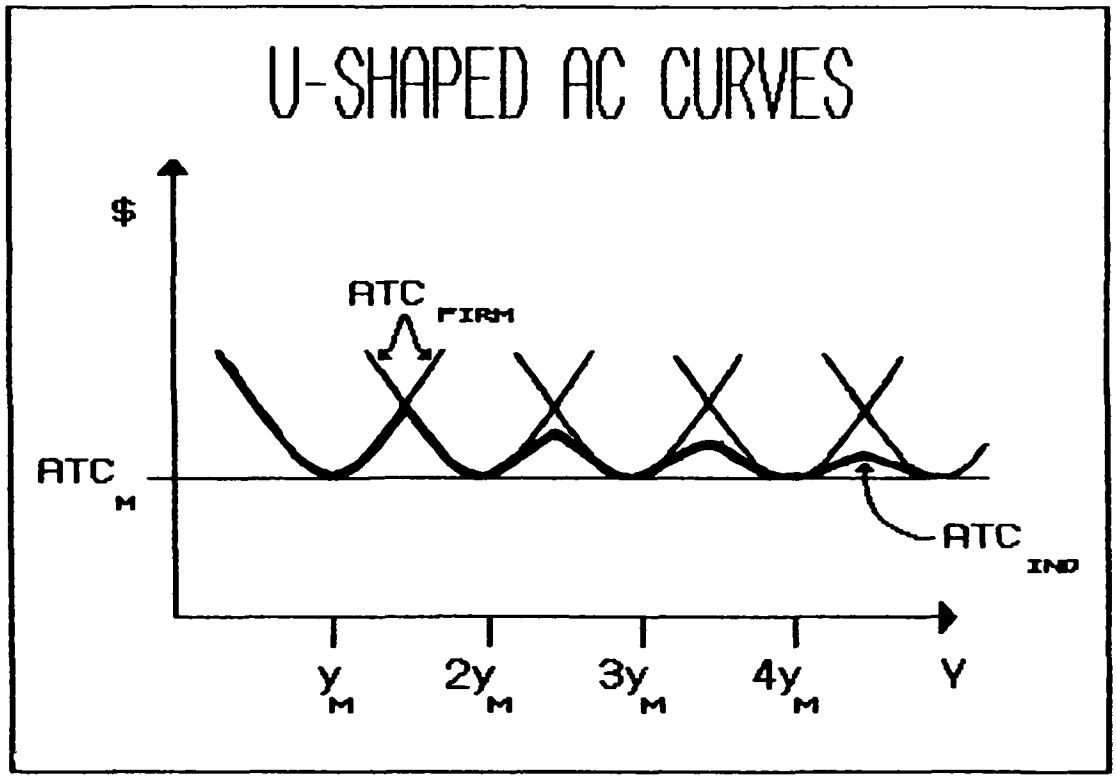


Fig 9. Industry Average Cost Curve
(Baumol et al., 1982a:32)

The industry average cost curve will reach minimum cost at output levels which are integer multiples of y_M, the average-cost minimizing output level for a single firm. The industry average cost can never be less than this value since no firm can produce at an average cost less than this. Also, at an integer multiple of y_M, say at ny_m , the industry average cost will clearly be minimized when n firms each produce an output of y_M. In such a configuration each firm is producing at a minimum point on the average cost curve,

so average cost must equal marginal cost. Thus, sustainable configurations clearly exist at integer multiples of y_M , since p = AC = MC (Baumol et al., 1982a:32).

Now consider a point between two of these minimums, say between y_M and $2y_M$. At output levels just above y_M single-firm production will still be the least costly way of supplying the output. As output levels continue to increase, at some point it will become cheaper for two firms to split the output (exactly where depends on the shape of the average cost curves). Up to this point the average cost of production will increase, and beyond this point (up to $2y_M$) it must decrease, because we have moved from the increasing cost side of a single firm's average cost curve to the decreasing cost side on two average cost curves. Similar behavior will be exhibited between any two minimums, and the industry average cost curve will have the scalloped shape shown. At any output level between two minimums, average cost will be either increasing or decreasing for the individual firms, so that cost does not equal marginal cost. Thus, no sustainable configurations can exist except at otuput levels which are integer multiples of y_M (Baumol et al., 1982a:32).

One final factor must be noted. As the output level becomes larger, the height of the peaks between the minimums decreases, eventually, in the limit as the number of firms becomes infinite, the peaks disappear totally. To see this, let C(y) denote the industry total cost function and c(y)

denote the firm total cost function. Any industry output, y, can be written as:

$$y = ny_M + d \tag{19}$$

where n is an integer and d is nonnegative and strictly less than y_M. Then:

$$C(y)/y = C(ny_M + d) / (ny_M + d)$$

$$\leq [C(ny_M) + C(d)] / (ny_M + d) \tag{20}$$

The inequality follows if it is assumed that the industry cost of producing an additional output of d is less than producing d by itself. So:

$$C(y)/y \leq [C(ny_M) + C(d)] / (ny_M + d) \tag{21}$$

$$\leq [nc(y_M) + C(d)] / (ny_M + d)$$

$$\leq \{ n [c(y_M) + (C(d)/n)] \}$$
$$/ \{ n [y_M + (d/n)] \}$$

$$C(y)/y \leq [c(y_M) + (C(d)/n)]$$
$$/ [y_M + (d/n)] \tag{22}$$

Taking the limit as n approaches infinity:

$$C(y)/y \leq c(y_M) / y_M \tag{23}$$

So, as the number of firms the market can support increases without bound, the average cost to the industry of any output equals the minimum short-run average cost to the firm. (It was already argued that the industry average cost cannot

be less than the minimum firm short-run average cost.)

Summary. Thus, with U-shaped short-run average cost curves, sustainable configurations exist only for output levels which are integer multiples of y_M, the output level which minimizes the firm's short-run average cost, or for output levels which can support a large number of firms in the market. Since every equilibrium configuration in a contestable market is a sustainable configuration (Proposition 4), these are the only output levels which would permit the market to be in equilibrium. The outlook, then, is not very optimistic.

Case II: Flat-Bottomed Average Cost Curves (Baumol et al., 1982a:32-37)

Now we will consider a slight, empirically justifiable modification to the previous situation. Istead of having a firm's average costs decrease to a minimum and then increase, consider the case in which the short-run average costs first decrease and then remain constant for a while before increasing. The average costs will then be minimized over a range of outputs rather than at a single output level. Figure 9 shows such a short-run average cost curve.

Let y_M denote the smallest output level at which the firm's short-run average costs are minimized. Also, assume

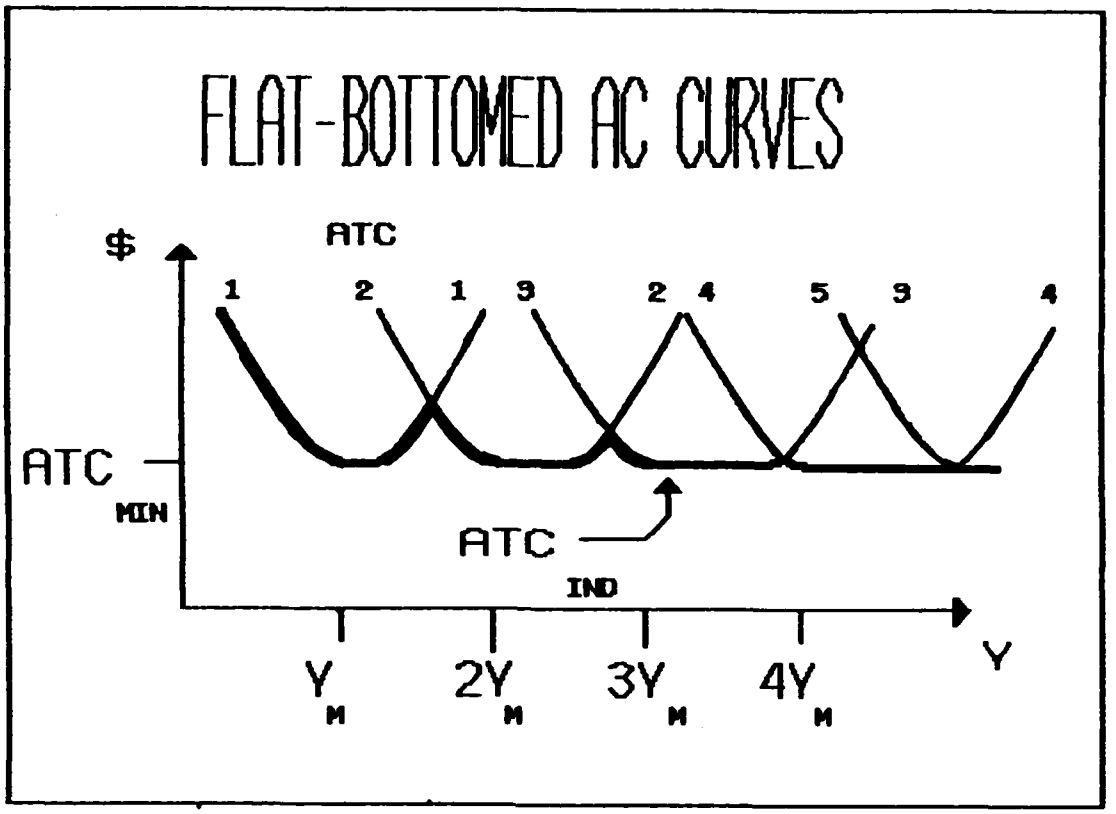


Fig 10. Industry Average Total Cost: Flat-Bottomed Curves

that the cost does not increase until output is greater than $(1 + k)y_M$. That is:

$$AC(y_M) = AC(y_1) = AC[\ (1 + k)y_M\ ]$$

$$\text{for all} \quad y_M \leq y_1 \leq (1 + k)y_M \qquad (24)$$

The industry average cost curve, shown in Figure 10, will assume the same scalloped shape as previously, except that the bottoms of the troughs will be flat. The region over which costs remain constant increases as the number of firms in the market increases. For now, assume that k is

less than unity. With only one firm costs remain constant over the region $[y_M , (1 + k)y_M]$, with two firms the constant cost region extends from $[2y_M , 2(1 + k)y_M]$ (since each firm produces over constant costs over $[y_M , (1 + k)y_M]$), for three firms it covers $[3y_M , 3(1 + k)y_M]$, and so on. The important point to notice is that the length of the constant cost interval increases from k, to $2k$, to $3k$, to nk when there are n firms in the market. Thus, with n firms in the market, where $nk = 1$, the industry can produce any output in range $[ny_M , n(1 + k)y_M]$ at constant cost. As n continues to increase, the product nk eventually will exceed unity. Then, the range of constant average cost is:

$$[ny_M , n(1 + k)y_M] = [ny_M , (n + nk)y_M]$$

$$\geq [ny_M , (n + 1)y_M] \qquad (25)$$

At this point, the range of constant cost for n firms extends up to (or beyond) the point at which the range of constant costs begin for $(n + 1)$ firms; the industry average cost curve becomes a horizontal line. Note that if $k = 1$, then the industry average cost curve is horizontal for all output levels above y_M.

Summary. So, with flat-bottomed average cost curves sustainable configurations can exist only where the industry average cost curve is horizontal. Thus, at low levels of output, that is, when there are few firms in the market, there may be output levels for which no sustainable configurations exist. There will always be an output level

such that a sustainable configuration will exist for every greater output. This minimum guaranteed sustainability output level varies inversely with the range of constant short-run average costs for the firm.

Multiproduct Markets

As noted before, the analysis of industry structure in multiproduct markets gets very complicated very quickly. However, some of the more basics concepts will be discussed here.

One of the basic concepts in multiproduct market analysis is a concept analogous to economies of scale. This is the concept of economies of scope. In a two product market, the goods exhibit economies of scope at a given output vector (y_1 , y_2) if the cost of producing the output vector (y_1 , y_2) is less than the cost of producing the output vectors $(y_1 , 0)$ and $(0 , y_2)$ (Baumol et al., 1982a:71-72). If there are economies of scope at y, then, since a sustainable configuration must minimize total industry cost, one firm (or several identical firms) must produce the output vector; that is, economies of scope are sufficient for the existence of multiproduct firms. Conversely, if multiproduct firms exist in a sustainable configuration, again they must minimize total industry cost. Thus, economies of scope must exist at y, so that economies of scope are also necessary for the existence of multi-product firms.

Very little can be said without going into a lot more detail. Beyond this point the mathematics becomes involved and beyond the scope of this effort. Those interested in more detail should consult Chapters Three through Eight of Baumol et al.'s book (1982a).

Bibliography

Allen, Robert F., "Shared Monopoly, Oligopoly, and Price-Cost Margins in American Manufacturing," (a) unpublished manuscript and (b) Referee's report. (undated).

Aurelius, Marcus, Meditations. Harmondsworth, England: Penguin Books, Ltd., 1964.

Bailey, Elizabeth and John Panzar. "The Contestability of Airline Markets During the Transition to Deregulation," Law and Contemporary Problems 44:125-145 Winter 1981).

Bain, Joseph S. Industrial Organization (second edition). New York: John Wiley and Sons, Inc., 1968.

-----. Barriers to New Competition. Cambridge: Harvard University Press, 1965.

Baumol, William J. et al. "Contestable Markets: An Uprising in the Theory of Industry Structure: Reply," The American Economic Review 73: 491-496 (June 1983).

----- et al. Contestable Markets and the Theory of Industry Structure. New York: Harcourt Brace Jovanovich, Inc., 1982a.

-----. "Contestable Markets: An Uprising in the Theory of Industry Structure," The American Economic Review 72: 1-15 (March 1982b).

Brock, William A. "Contestable Markets and the Theory of Industry Structure: A Review Article," Journal of Political Economy, 91: 1055-1066 (November/December 1983).

Coursey, Don, et al. "Market Contestability in the Presence of Sunk (Entry) Costs," The Rand Journal of Economics 15: 69-84 (Spring 1984a).

----- et al. "Natural Monopoly and Contested Markets: Some Experimental Results," The Journal of Law and Economics 27: 91-113 (April 1984b).

Demsetz, Harold. "Why Regulate Public Utilities?," The Journal of Law and Economics 2: 55-69 (April 1968).

Dixit, Avinash. "Recent Developments in Oligopoly Theory," The American Economic Review Proceedings 72: 12-17 (May 1982).

Harrison, Glenn and Michael McKee. "Monopoly Behavior, Decentralized Regulation, and Contestable Markets: an Experimental Evaluation," The Rand Journal of Economics 16: 51-69 (Spring 1985).

MacDonald, James M. "Entry and Exit on the Competitive Fringe," Southern Economic Journal 52: 640-652 (January 1986).

Needham, Douglas. The Economics of Idustrial Structure, Conduct, and Performance. New York: Saint Martin's Press, 1978.

Nicholson, Walter. Microeconomic Theory: Basic Principles and Extensions (Second Edition). Minsdale, Illinois: The Dryden Press, 1978.

Schwartz, Marius and Robert J. Reynolds, "Contestable Markets: An Uprising in the Theory of Industry Structure: Comment," The American Economic Review 73:488-490 (June 1983).

Shepherd, William D. "'Contestability' vs. Competition," The American Economic Review 74: 572-587 (September 1984).

Spence, Michael. "Contestable Markets and the Theory of Industry Structure: A Review Article," Journal of Economic Literature 21: 981-990 (September 1983).

Stigler, George J. The Organization of Industry. Homewood, Illinois: Richard D. Irwin, Inc., 1968.

VITA

Second Lieutenant Patrick M. Lundy was born on 12 August 1963 in Chicago, Illinois. He graduated from high school in Mundelein, Illinois, in 1981 and attended the University of Michigan from which he received the degree of Bachelor of Science with concentrations in Math and Physics in 1985. Upon graduation, he received a commission in the USAF through ROTC. He immediately entered the School of Engineering, Air Force Institute of Technology, in May of 1985.

Permanent address: 873 Liberty Bell Lane
Libertyville, Illinois 60048

FND

5-87

DTIC